

PACIFIC LINGUISTICS

*Series C - No. 49*

SOUTH-EAST ASIAN LINGUISTIC STUDIES

VOL. 4

Nguyen Dang Liem, ed.



Department of Linguistics

Research School of Pacific Studies

THE AUSTRALIAN NATIONAL UNIVERSITY

PACIFIC LINGUISTICS is issued through the *Linguistic Circle of Canberra* and consists of four series:

- SERIES A - OCCASIONAL PAPERS
- SERIES B - MONOGRAPHS
- SERIES C - BOOKS
- SERIES D - SPECIAL PUBLICATIONS

EDITOR: S.A. Wurm.

ASSOCIATE EDITORS: D.C. Laycock, C.L. Voorhoeve, D.T. Tryon, T.E. Dutton.

EDITORIAL ADVISERS:

B. Bender, University of Hawaii  
D. Bradley, University of Melbourne  
A. Capell, University of Sydney  
S. Elbert, University of Hawaii  
K. Franklin, Summer Institute of  
Linguistics  
W.W. Glover, Summer Institute of  
Linguistics  
G. Grace, University of Hawaii  
M.A.K. Halliday, University of  
Sydney  
A. Healey, Summer Institute of  
Linguistics  
L. Hercus, Australian National  
University  
N.D. Liem, University of Hawaii

J. Lynch, University of Papua  
New Guinea  
K.A. McElhanon, University of Texas  
H. McKaughan, University of Hawaii  
P. Mühlhäusler, Linacre College,  
Oxford  
G.N. O'Grady, University of  
Victoria, B.C.  
A.K. Pawley, University of Hawaii  
K. Pike, University of Michigan;  
Summer Institute of Linguistics  
E.C. Polomé, University of Texas  
G. Sankoff, Université de Montréal  
E. Uhlenbeck, University of Leiden  
J.W.M. Verhaar, University of  
Indonesia, Jakarta

ALL CORRESPONDENCE concerning PACIFIC LINGUISTICS, including orders and subscriptions, should be addressed to:

The Secretary,  
PACIFIC LINGUISTICS,  
Department of Linguistics,  
School of Pacific Studies,  
The Australian National University,

Canberra, A.C.T. 2600.  
Australia.

Copyright © The Authors.  
First published 1979.

The editors are indebted to the Australian National University for help in the production of this series.

This publication was made possible by an initial grant from the Hunter Douglas Fund.

National Library of Australia Card Number and ISBN 0 85883 201 1

## TABLE OF CONTENTS

	<i>Page</i>
COMPARATIVE ASPECTS OF LUE SYNTAX, by John F. Hartmann	1-15
CASES AND CLAUSES IN LAO, by Mary E. Honts	17-38
A CLASSIFICATION OF THAI CLASSIFIERS, by SamAng Hiranburana	39-54
HOW MUCH IS ENGLISH INFLUENCING THE LANGUAGE OF THE EDUCATED BANGKOK THAIS?, by Wilaiwan Kanittanan	55-59
FORMS AND MEANINGS OF THE THAI PARTICLE <u>SI</u> , by Joseph R. Cooke	61-98
PHONOLOGICAL REDUCTION OF SOME FINAL PARTICLES IN MODERN THAI, by Patcharin Peyasantiwong	99-115
HOW MANY TONES FOR SOUTHERN THAI?, by Anthony Diller	117-129
TONAL RULES FOR ENGLISH LOANWORDS IN THAI, by Jack Gandour	131-144
COMPLICATIONS IN TEMPORAL PREVERBS AND THEIR SEMANTIC INTERPRETATION, by Udom Warotamasikkhadit	145-153
LES PRONOMS PERSONNELS DU KHMER: ORIGINE ET EVOLUTION, by Saveros Pou	155-178
STANDARDISATION AND PURIFICATION: A LOOK AT LANGUAGE PLANNING IN VIETNAM, by Nguyen Dinh-Hoa	179-205
A STATISTICAL COMPARISON OF VERBS AND NOUNS IN RŒGLAI, by Maxwell Cobbey	207-212

	<i>Page</i>
TWO PROGRESSIVE ASPECT MARKERS IN CHINESE, by Anna Kwan-Terry	213-232
A NOTE ON THE TAGALOG PASSIVE IN THE TOTANES MANUSCRIPT, by Joseph F. Kess	233-239
KELEY-I PHONOLOGY AND MORPHOPHONEMICS, by Lou Hohulin and Michael Kenstowicz	241-254
SOME FEATURES OF NARRATIVE DISCOURSE IN KADAZAN, by Hope M. Hurlbut	255-282
THE VERBAL PARTICLE <u>LEU</u> IN THE MAUMERE LANGUAGE, by Joan M. Rosen	283-303
THE USE OF DEICTICS TO IDENTIFY PARTICIPANTS IN BAHASA INDONESIA, by Donald F. Barr	305-313
QUANTIFIERS IN JAVANESE AND INDONESIAN, by Marmo Soemarmo	315-363
VOWEL PATTERNING AND MEANING IN MALAY PAIR-WORDS, by Tham Seong Chee	365-377
EXPRESSIVES IN KEDAH MALAY, by James T. Collins	379-406
THE SAMOAN CONNECTION, OR: VERB AND NOUN PHRASE RELATIONS IN A POLYNESIAN LANGUAGE, by Michael R. Thomas	407-413
PERCEPTUAL DIMENSIONS OF CANTONESE TONES: A MULTIDIMENSIONAL SCALING REANALYSIS OF FOK'S TONE CONFUSION DATA, by Jack Gandour	415-429
REVIEW ARTICLE: ROBERT K. HEADLEY JR, KYLIN CHHOR, LAM KHENG LIM, LIM HAK KHEANG, CHEN CHUN: CAMBODIAN-ENGLISH DICTIONARY, by Philip N. Jenner	431-436

*EDITORIAL NOTE:* Most of the articles included in this volume were received in 1976 to 1978, and theoretical views expressed in them may be at variance with the authors' present views on the subject. This note has been inserted at the specific request of the Editor for this volume on behalf of the authors.



## COMPARATIVE ASPECTS OF LUE SYNTAX\*

JOHN F. HARTMANN

Studies of comparative Tai syntax are rare aside from a few studies of pronouns and classifiers. There is the rather complacent belief that there are few dramatic differences in syntax between dialects. The thesis of this article is that if we consider pragmatics along with semantic-syntactic structures, important distinctions are found. The differences are subtle ones, but they have important implications for linguistic theory and our understanding of Tai dialects in general.

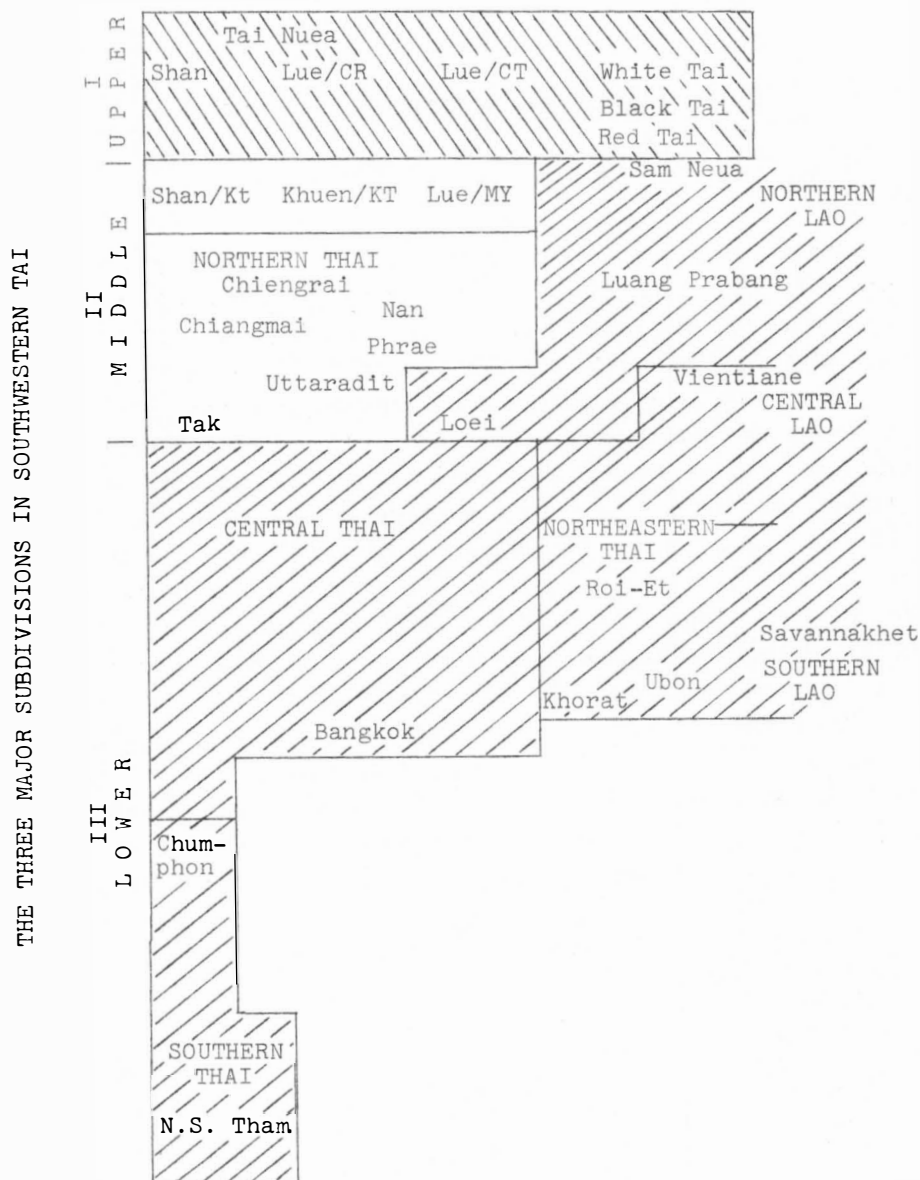
Tai-Lue, the focus of this analysis, is a Tai dialect whose centre is at Chieng Hung, in a region called Sipsong Panna in the southwestern part of Yunnan, China. It is a dialect whose phonological features overlap with neighbouring Shan to the west in Burma, with Northern Thai directly south in Thailand, and with White Tai spoken in adjacent areas of Laos and North Vietnam.

Still another dialect of Lue is spoken at Moeng Yong, Burma. In terms of tonal splits and minor phonological features, it is identical to Tai-Khuen of Kentung, Burma and the Thai dialects of Northern Thailand. For a clear picture of these relationships, a suggested alignment of dialects in Southwestern Tai appears on the following page.

---

\* From Chapter V of my dissertation: *The Linguistic and Memory Structure of Tai-Lue Oral Narrative*. (1976) Reproduced by University Microfilms, Ann Arbor.

## SUGGESTED ALIGNMENT OF DIALECTS IN SOUTHWESTERN TAI



The general shape of the tones of Chieng Rung are shown in the following diagram using both a system of numbered tones and their description in words along with the scheme used often in phonological description of Chinese tones. In the latter system, a pitch level of 5 is high and 1 is low; 3 would be in the mid range.

	*A	*B	*C	*D-long	*D-short
*vl (yin)	1 high-level 7 55	2 mid-rising 1 35	3 low, glot., slt. rise J 13	=2	=1
*vd (yang)	4 fall-ing V	5 mid-level ┐	6 low, level, slt. rise └	=5	=5
	Smooth Syllables			Checked Syllables	

In this matrix, the maximum number of six tones are found on the smooth syllables. The tones on the checked syllables are matched up with their nearest counterparts in the class of smooth syllables. As explained elsewhere, the tones of the smooth and checked are conditioned by different variables so that they stand in complementary distribution. The tones shown here, then, are phonemic, not phonetic. In some works, e.g. Purnell (1963), it is not always clear whether the tones which are enumerated are phonemic or not. Closer examination reveals that the seven tones of Northern Thai, for example, are phonetic; only six can be isolated on free syllables.

The data on which the comparative study of Tai syntax comes is based on a variety of sources. One is my own study of orally composed chanted narratives called /khap<sup>1</sup> lɔ̌<sup>6</sup>/ 'to sing in the Lue manner'. This material, referred to later, is that of two male singers: an older man of about 60 (Text I) and a younger man of about 40 (Text II). Otherwise the data are from my field notes or those of others whose names are cited.

Particular attention will be paid to utterance final particles. Definite underlying semantic-syntactic differences exist and form a marked communication boundary between dialects despite innocent-looking minor surface changes in lexical shapes.

The negative and the negative interrogative (question particles) are the most outstanding illustrations of special aspects of Lue syntax that shall be dealt with. In addition, comment will be made concerning the semantic-syntactic contrast involving word change between Siamese and Lue use of 'can' vs. 'to be able to'. Finally, discourse level syntactic functioning of pronouns and particles which punctuate clause and paragraph divisions in the Lue narrative will be discussed.

#### A. Interrogative forms

Lue questions and related responses entail presuppositions that do not exactly parallel either Siamese or Northern Thai usage. The word order is, for the most part, the same: question particles are utterance final. It is best to examine some of the Lue rules on their own terms before making any comparisons with other dialects.

##### (1) -aa<sup>5</sup>, -aa<sup>4</sup>

The first particle, -aa<sup>5</sup>, is used in interrogative utterances that call for information, i.e. the usual yes-no type of question. It is used in structures that do not have other question words such as wh-forms: 'what, where, why, how', etc. Where the Lue equivalent of the English wh-forms appears, the tone of the question particle changes from tone 5 (mid level) to tone 4 (mid falling). Some examples are:

- (a) dii<sup>1</sup> -aa<sup>5</sup>  
good Q-Pt.

'Is it good?'

- (b) pin<sup>1</sup> kun<sup>4</sup> tii<sup>5</sup> nay<sup>1</sup> -aa<sup>4</sup>  
be person place where Q-Pt.

'Where are you from?'

##### (2) -aa<sup>5</sup> vs. kaa<sup>4</sup>

The final question particle, explained above, contrasts with kaa<sup>4</sup> in that the latter is used in questions with an underlying presupposition: 'I assume that it is the case that', or 'right?', as glossed in the example given below. The particle kaa<sup>4</sup> is used both in the initiating question and in the expected response. The underlying presupposition can be confirmed or refuted with an affirmative or negative response. In its confirmative function, kaa<sup>4</sup> has the force of a mildly emphatic particle. The following examples are illustrative but not completely so. More data are needed.

- (a) kin<sup>1</sup> kaa<sup>4</sup> (Question)  
eat, right?

'(someone) eat, right?'

- (b) kin<sup>1</sup> kaa<sup>4</sup> (Response)  
 eat, right!  
 'Right, (someone) eats'

It is noteworthy that the constructions used for asking and for answering shown above are synonymous. Without some notion of pragmatics, performatives, or context of situation, their semantic difference could not be understood.

In Northern Thai (Purnell 1963) we find the same form and function: kää, kàa. A further distinction is made between the former (mid tone) affirmative particle and the latter (low tone) emphatic particle. The interrogative function appears to parallel the use of Siamese rɯ̌ in some contexts and chây máy in others. The mildly emphatic kaa<sup>4</sup> shared by Lue and Northern Thai mirrors the Siamese form sì.

On the other hand, the Lue question particles -aa<sup>5</sup>, -aa<sup>4</sup> do not appear to have a reflex in Northern Thai or Siamese. The Northern Thai parallel appears to be kǎo, which is roughly equivalent to Siamese máy. Northern Thai is still different from both Lue and Siamese in having bǎo, the question particle that has as its underlying presupposition the paraphrase: 'may I invite you to ...', as in, for example:

- |      |                          |     |                   |
|------|--------------------------|-----|-------------------|
| N.T. | kin bǎo                  | vs. | kin kǎo           |
|      | eat Q-Pt.                |     | eat Q-Pt.         |
|      | 'Would you like to eat?' |     | 'Are you eating?' |

It seems that even in Northern Thai, the distinction between the invitational interrogative bǎo and the informational interrogative kǎo is disappearing in favour of the former.

To return to Lue, in place of kaa<sup>4</sup> as a response particle, which might be described as a simple affirmative particle, we find the more strongly emphatic particle yaa<sup>2</sup>. At the other extreme, the most neutral particle, used simply to punctuate an utterance, is the form ɿɛ<sup>5</sup>. It is very frequently used in the Lue oral narrative, especially in the performance of the second singer. The following examples show the contrasting function between the sharply emphatic and the emotionless punctuating utterance final particles in Lue.

- (a) yuu<sup>2</sup> kaa<sup>4</sup>  
 here right?  
 '(someone) is here, right?'
- (b) yuu<sup>2</sup> yaa<sup>2</sup>  
 here Pt.-emph. affirm.  
 '(someone) is here, indeed!'
- (c) yuu<sup>2</sup> ɿɛ<sup>5</sup>  
 here Pt.-punct. affirm.  
 '(someone) is here.'

As recent debate concerning a theory of speech acts attests, there are many difficulties to be encountered in assigning an underlying performative or in formulating the most precise sets of presuppositions to the utterance we have been discussing. The case for Lue and other Tai dialects is facilitated, however, by the presence of contrasting particles. Matisoff (1973) has made the claim that particles are degenerate verbs. Further support for a verbal analysis, and even an underlying clause of presupposed information, comes from the work of Day (1966), who assigns separate clause rank (level in tagmemics) to utterance final particles in Tho, a Tai language of North Vietnam. The difficult semantics of utterance final particles remains however. The commonplace observation is that they parallel the use of intonation and stress in English. Noss (1964) rightly notes that sentence particles indicate speaker attitude, whose meaning "can be only vaguely stated, because a great deal depends upon the emotional interplay between speakers." A complete analysis of particles in Lue would be a major undertaking calling for more natural, conversational data than are found in our corpus of oral narratives.

It should be noted that Ross (1970) uses Thai (Siamese or Central Thai) in arguing for a performative verb and I-you axis in the deep structure of declarative sentences. While his basic argument is sound, his information on Thai is faulty. He states, incorrectly: "In this language, every sentence must end with the particle *kháp* or *khâ*."

Such is not the case. The appearance or non-appearance of the utterance final 'polite' particles *kháp* (masculine) and *khâ* (feminine), are optional to begin with. From what I have seen, the former is limited to Siamese and Northern Thai, the latter to Siamese. More important is that their occurrence is dependent on several interrelated contextual factors. First is the status of the speaker and hearer; second is the emotional force between them. The first parameter concerns social distance, the second psychological (phatic communion possibly). More specifically, an adult (age status) would normally never use *kháp* or *khâ* in speaking to a child or other persons of inferior social status. Likewise, when there are no constraints calling for verbal displays of deference or politeness, the 'polite particle' is not used, or when other negative emotional states would overrule its use, such as anger.

Ross uses the final particles *kháp* and *khâ* as evidence for an underlying 'I'. He labels the Thai particles "utterer agreement particles (UAP)." As indicated in the preceding discussion, statements from Day, Noss, and Matisoff can be used to show that many Thai particles are manifestations of performative verbs which have been weakened rather than deleted following Ross's rule for "Performative Deletion". Following

Noss's argument that sentence-final particles are an indicator of speaker's attitude, they would carry a perlocutionary force and hence should be classified as performative verbs in many cases. The case for particles is not always that clear, however.

In his study of Central Thai syntax, Scovel (1970) shares the opinion that the historical origins of particles must be studied before they can be understood completely. My own reaction is that the issue must be decided on synchronic evidence. Nevertheless, we can point to some limited historical information that might inspire others to make a thorough study of older historical texts.

To further complicate the picture of the 'polite' particle used by today's Siamese women, we note that there are three tonal forms, already an indication that something more complex than an 'utterer agreement particle' is involved. As we examine the three tonal shapes, the case for particles-as-performative verbs is strengthened.

#### Siamese (Central Thai):

##### A. Urban-Refined speech (being sophisticated)/Formal

- (1) mɛɛ khǎa  
mother Pt.-female-endearment-intimate: to call sm.  
'Mother, dear' (call for attention)
- (2) mii máy khá  
have Q-Pt. Pt.-female-deferential  
'Do you have any?' (I DEFER to you)
- (3) mây mii khâ  
neg. have Pt.-female deferential  
'I don't have any' (I DEFER to you)

##### B. Rural, polite speech (being nice)/Informal

- (1) mɛɛ cǎa
- (2) mii máy cá
- (3) mây mii câ

From the above, we see the further division between urban-sophisticated form and rural-'nice' particles. The two overlap, depending on social setting. In rural speech, the basis of Thai society, it would be rare indeed to hear anyone use the first set of formal 'polite' particles. On the other hand, the second set of particles would often be used in an urban setting or one of less formal demands, the market for example. Also, in an urban context, a 'superior' would use the second set in addressing an 'inferior'.

As for the historical origins of the polite particles, one must rely on older historical texts. In the plays of Rama VI, a brief glance reveals the following. Both men and women used the particle *khâ*, which is limited to female use today. Moreover, the full older form seems to be *câw khâa* 'my lord'. (The form *khâa* is glossed as 'slave, I'.)

On still another level indicating probable Cambodian origins, we find in the 'Royal vocabulary' the forms *phâyâ khâ*, the utterance final polite form used by men and *phêe khâ*, used by women. Both are used in addressing the King and Queen, but not the reverse.

While the female particle *khâ* might have its origins in the noun *khâa* 'slave', the male particle appears to have verbal origins. It is believed that the base is *khăo ráp* 'ask to receive'. The fact that two different likely historical sources, a noun and a verb, are indicated for the female and male polite particles confuses rather than clarifies the issue concerning their synchronic status: UAP or performative verb.

In other Tai dialects, the tendency is to use a final unisex 1st or 2nd person formal pronoun.

Lue:	<i>khoy</i> <sup>3</sup>	'I', m./fem. (inferior to sup.); polite single-word response
Lao:	<i>khan</i> <sup>3</sup> <i>oy</i>	'I' (inferior to superior); polite response word (from Roffe)
Northern Thai:	<i>căw</i>	'yes'; polite word (fem.); 'you'
	<i>khăp</i>	'yes'; man's polite word (from Purnell)

At this point, after considering both diachronic and synchronic evidence from four Tai dialects, one might conclude that utterance final particles are evidence for an underlying 'I' in a deep structure performative clause, i.e. the Ross argument. But as already pointed out, the 'polite particles' are only one set in a catalogue of many other particles which have verbal form in many cases and performative function on all occasions. It may be that the Performative Deletion rule has to be amended to read that, in the case of Tai dialects, sometimes all or only part of the whole clause is deleted leaving either an NP-subject (*khâ*) verb (*kháp*) or NP-object (*căw* 'you' Northern Thai). In any case, the combined insights of Day, Matisoff and Noss indicate that utterance final particles (one or more) are manifestations of a performative clause.

The fact remains that the presence of utterance final particles in Tai dialects and many other South-East Asian languages points to the incompletely analysed inter-relatedness of pragmatics, syntax and semantics. A broader perspective on the pragmatics of 'linguistic etiquette' (e.g. Tai particles) that seems applicable to all of South-East Asia is



found in Geertz's (1960) statement, which we use as a fitting conclusion to this segment of the discussion.

It has already been pointed out how etiquette patterns, including language, tend to be regarded by the Javanese as a kind of emotional capital which may be invested in putting others at ease. Politeness is something one directs toward others; one surrounds the other with a wall of behavioral formality which protects the stability of his inner life. Etiquette is a wall built around one's inner feelings, but it is, paradoxically, always a wall someone else builds, at least in part. He may choose to build such a wall for one or two reasons. He and the other person are at least approximate status equals and not intimate friends; and so he responds to the other's politeness to him with an equal politeness. Or the other is clearly superior, in which case he will, in deference to the other's greater spiritual refinement, build him a wall without any demand or expectation that you reciprocate.

#### B. Negative Interrogative

We begin this part of the discussion with contrasting examples from Siamese and Lue.

S: kin rĩi plaàw eat or neg. 'Did you eat (it) or not?'	Lue: kin <sup>1</sup> m̃ <sup>2</sup> kin <sup>1</sup> eat neg. eat 'Did you eat (it) or not?'
---	--

The interesting features lie in the Siamese constructions. The Siamese rĩi by itself can function as a question particle that has as its underlying presupposition, 'I assume that you', as in:

Siamese: kin rĩi  
eat Q-Pt.  
'(somebody) eats/ate, I assume'

Likewise, the Siamese form plaàw can be used as a single-word response which rejects the questioners presupposition. The Lue negative m̃ (or its full form baw<sup>2</sup>) cannot be used alone in a response. The fact that the Siamese forms khráp, khâ, rĩi and plaàw function as single-word responses strengthens the argument for performative verb status for utterance final particles.

Still other differences can be found which show striking dissimilarities between Siamese and Lue syntax.

S: kin lɛ́w rĩi yaŋ eat already or yet 'Did you eat yet?'	Lue: kin <sup>1</sup> lɛw <sup>6</sup> laa <sup>5</sup> (or laa <sup>2</sup> ) eat already Q-Pt. 'Did you eat yet?'
---	---

While Siamese lɛ́w and Lue lɛw<sup>6</sup> 'already' point to common lexical and syntactic origins, the appearance of rĩi and yaŋ in Siamese shows a divergence. Two other examples contrast Siamese rĩi with Lue di?<sup>1</sup> 'or'.

S: ca? kin náam plaàw rɿɿ náam chaa  
 will eat water plain or water tea  
 'Will you drink water or tea?'

L: di?¹ kin¹ nam⁶ kat¹ / di?¹ kin¹ nam⁶ laa⁶ -aa⁵  
 will eat water cool will eat water tea Q-Pt.  
 'Will you drink water or tea?'

Noticeable in the Lue citation is the absence of a conjoining device aside from pause at the syntactic boundaries between clauses.

Part of the explanation for the syntactic differences between Siamese and Lue, especially with the uniqueness of rɿɿ in Siamese, might come from possible borrowing from Cambodian. Huffman (1973) claims that the borrowing has been in the reverse, from Thai into Cambodian. Claims for directionality aside, Siamese and Cambodian do share the following forms which do not appear in Lue. The citations are from Huffman.

Thai	Cambodian	
rɿy	rɿɿ	'Q. pt in either/or Q's.'
lɛ́w-rɿy-jan	haay-rɿɿ-nɿw	'already or not-yet'
laay	laay	'at-all'
ná?	nah	'final-hortatory-particle'

Another area in which the negative and interrogative interact is with the form ɱ saŋ¹, which may be a case of idiomatic usage. We find

L: ɱ pin¹ saŋ¹  
 neg. be what/thing  
 'Never mind'

S: mây pen ray  
 neg. be what/thing  
 'Never mind'

L: kin¹ ɱ-saŋ¹  
 eat what  
 'What are you eating?'

S: kin ʔaray  
 eat what  
 'What are you eating?'

The negative does not appear in the Siamese form for 'what are you eating?'. The question is why the syllabic ɱ appears in the affirmative interrogative kin¹ ɱ saŋ¹. I suspect that the appearance of the syllabic ɱ in this latter case is a purely phonetic phenomenon with no syntactic or semantic relevance. An alternate form is kin¹ ʔii²-saŋ¹. Further evidence of the idiosyncratic nature of ɱ is that we find in Chiengkham Lue (Hartmann 1975) two similar variations for the vocative ɱ¹ pɔɔ⁵, ʔii² pɔɔ⁵ 'father dear'. In Northern Thai (Purnell 1963) we

find the same phonetic process of the intrusion of  $\eta$  in  $\eta$ impɔɔ 'father dear',  $\eta$ imɛɛ 'mother dear'. Further investigation would undoubtedly shed more light on the peculiar grammar of  $\eta$  saŋ<sup>1</sup>.

Some additional examples of Lue forms related to the foregoing discussion are the following ones taken from Gedney's unpublished field notes. The abbreviation R and Y stand for Lue of Chieng Rung and Moeng Yong respectively.

- (1) R yuu<sup>2</sup> baw<sup>2</sup> yuu<sup>2</sup>  
 Y yuu<sup>2</sup> -aa<sup>2</sup>  $\eta$  yuu<sup>2</sup>-aa<sup>2</sup>  
 stay Q. neg. stay Q.  
 'Is he here or not?'
- (2) Y & R  $\eta$ <sup>2</sup> dɔy<sup>3</sup> tɛɛ<sup>5</sup>  
 neg.eat pt.  
 'Haven't eaten yet'
- (3) Y & R dɔy<sup>3</sup> yaŋ<sup>4</sup>  $\eta$ <sup>2</sup> lɛw<sup>6</sup> tɛɛ<sup>5</sup>  
 eat yet neg. already pt.  
 'Haven't finished eating'

#### C. day<sup>3</sup> vs. caŋ<sup>1,5</sup> and change of word order

Here we shall point to a minor transformation. The Lue form day<sup>3</sup> 'can' has the Siamese reflex dây. On the other hand, caŋ<sup>1</sup> is synonymous with Siamese pen 'to be able, to have the skill'. However, the latter form results in a change in word order.

- S: kin dây máy  
 eat can Q-Pt.  
 'Can you eat it?'
- L: di?<sup>1</sup> kin<sup>1</sup> day<sup>3</sup> -aa<sup>5</sup>  
 will eat can Q-Pt.  
 'Can you eat it?'
- S: kin pen máy  
 eat able Q-pt.  
 'Do you know how to eat it?'
- L: caŋ<sup>1</sup> kin<sup>1</sup> -aa<sup>5</sup>  
 able eat Q-pt.  
 'Do you know how to eat it?'

The grammar of the Lue form day<sup>3</sup> is different too in that it calls for the future particle di?<sup>1</sup> as a pre-verb. The Lue form caŋ<sup>1</sup> is shared with Northern Thai.

#### D. Pronouns

As with the particles, a grammar of pronominal usage depends on contextual factors, especially those dealing with social and situational dimensions. The first person singular is the same in Lue as in Lao, another indication of shifting cultural affinities among Lue, Shan, Lao, and Northern Thai. The form *khoy*<sup>3</sup> 'I' would be used in a formal social setting. It seems, on the basis of our limited conversational data, that the same form, *khoy*<sup>3</sup>, has the double function of final 'polite' or single-word response particle: 'yes, politely'. On a still higher level of formality, we find *khoy*<sup>3</sup> *bat*<sup>1</sup> 'I'.

On the intimate level, *kuu*<sup>5</sup> 'I' and *m+ŋ*<sup>4</sup> 'you' are used by men and women alike.

In the semi-formal context of family life and celebrations kin terms are used. Such is the case in the Lue chant (*khap*<sup>1</sup> *l++*<sup>6</sup>). The singers use kin terms and a *nom-de-plume* in referring to themselves - usually at the beginning or ending of their participation in the event.

In Text I, the singer is the senior member and refers to himself as *po*<sup>5</sup> 'father', *pi*<sup>5</sup> 'elder sibling' or his *nom-de-plume* *sii*<sup>1</sup> *taa*<sup>1</sup> *dam*<sup>1</sup> 'black eyes'. In Text II, the singer is a generation younger, and although a man in his late thirties, he refers to himself as *luk*<sup>5</sup> 'child, son' in the presence of the older singer. But he uses the non-kin term *pan*<sup>5</sup> 'I' that a man would use in addressing a younger woman such as his wife, whose implied social status is inferior to his own. In this particular use of *pan*<sup>5</sup> 'I' (also 'they') the younger man refers to a hypothetical woman, the female co-singer that is usually used in the chanting of the Lue tales. Such a performance, and the pronominal system, has an underlying structure of male-female dialogue (cf. Klammer 1973 for his discussion of dialogue as the basic unit of discourse).

#### E. Discourse Level Particles

Of the catalogue of particles that are found in Lue, a few deserve added comment with respect to their role in the syntax of the chanted narrative. In Text II, the younger singer's, several particles are used as oral punctuation and a stylistic device as well. Many of his lines are punctuated by a final *nii*<sup>6</sup>, *ni*<sup>-6</sup> or *le?*<sup>5</sup>. Both are lexical manifestations of a syntactic and phonological boundary. In the actual performance, this is not at all obvious to the outsider. There are no necessary phonological breaks or pauses in the course of singing. But enough cues are given to assert the existence of co-occurrent phonological, grammatical and lexical boundaries. When they do not co-incide exactly, they may be said to overlap. From the standpoint of both method and theory this is very important. In the process of analysing one's

data, cues from these three intertwining fields must be sought in the course of segmenting an otherwise continuous stream of chanted speech. My first reaction on hearing these semantically empty particles, particularly  $le?^5$ , was to remove them from the data as irrelevant. Taking the lead from the tagmemic framework, I discovered that in an oral grammar, they serve the important function of audible punctuation marking sentence boundaries or clause divisions. On a higher level, clause groups, we find phrase-length particles used to punctuate larger units:  $duu^1 le?^5 cam^1$  and  $bat^1 dew^1 van^4 nii^6$ .

We conclude this paper on Lue syntax with a final comment on the place of particles in linguistic theory. Some sort of a 'performative clause' analysis as seen in the earlier work of Ross (1970) is called for. He has argued that a performative solution is preferable to a pragmatics because the former allows for greater formality. Its mechanics are much neater, for one thing. After going to great lengths to prove the existence of a universal underlying performative clause, the suggestion seems to be that the resulting structure is more real than a pragmatics solution. A pragmatics, by comparison, is less real because it rests too many of its claims on things, to quote Ross, "in the air".

Tagmemics and Firthian theory could assume as axiomatic the presence of a speaker-hearer or I-Thou structure in the social situation and the very act of communication, the major function of language. This dialogic structure would be neither "in the air" nor in a syntactic deep structure. It is considered an empirically observable fact which need not be intuited from or argued out of the data of isolated utterances.

BIBLIOGRAPHY

DAY, A.C.

- 1966      The Syntax of Tho A Tai Language of Vietnam.    Ph.D.  
             dissertation, University of London.

GEERTZ, Clifford J.

- 1960      *The Religion of Java*.    Glencoe, Ill.: The Free Press.

HARTMANN, John F.

- 1975      'Syllabic ɱ in Tai-Lue and Neighboring Tai Dialects'. In:  
             Thomas W. Gething, ed. *Tai Festschrift for William J.*  
             *Gedney. Southeast Asian Studies Working Paper No.8.*  
             Honolulu: University of Hawaii. Also published as Gething  
             and Nguyen Dang Liem, eds. 1979 *Papers in South-East Asian*  
             *Linguistics*, No.6: *Tai Studies in Honour of William J.*  
             *Gedney*. PL, A-52.

HUFFMAN, Franklin E.

- 1973      'Thai and Cambodian - A Case of Syntactic Borrowing?'.  
             *Journal of the American Oriental Society* 93:88-509.

KLAMMER, Thomas

- 1973      'Foundations for a Theory of Dialogue Structure'. *Poetics*  
             9:27-64.

MATISOFF, James A.

- 1973      Class Lecture Notes.    Linguistics Institute, University  
             of Michigan, Ann Arbor.

NOSS, Richard B.

- 1964      *Thai Reference Grammar*. Washington, D.C.: Foreign Service  
Institute.

PURNELL, Herbert C.

- 1963      *A Short Northern Thai-English Dictionary*. Chiangmai:  
Overseas Missionary Fellowship.

ROSS, Robert J.

- 1970      'On Declarative Sentences'. In: Roderick A. Jacobs and  
P.S. Rosenbaum, eds *Readings in English Transformational  
Grammar*, 222-72. Waltham, Mass.: Ginn.

SCOVEL, Thomas S.

- 1970      *A Grammar of Time in Thai*. Ph.D. dissertation, University  
of Michigan, Ann Arbor.





## CASES AND CLAUSES IN LAO

MARY E. HONTIS

The purpose of this paper is to present a syntactic analysis of Lao at the clause level. The format and approach employed will be that used by Nguyen Dang Liem in the study *Cases, Clauses and Sentences in Vietnamese* (Liem 1975). This approach utilises the principles and methods of both tagmemics and the lexicase model of case grammar. This eclectic approach has been developed in this work and in earlier studies (Cook 1971, Liem 1971a, 1971b, 1974 and 1976, Heidi Platt 1970, and John T. Platt 1971) to provide a suitable framework for studies in contrastive analysis. Although the theoretical orientations of tagmemics and those of case grammar recognise and attempt to describe models of linguistic structure fundamentally different in many respects, the combination of their basic principles employed in this framework can provide a systematic and pedagogically useful analysis of some of the basic syntactic and semantic properties of a language for contrastive purposes. The findings of this study will be presented in two sections, the first dealing with case relations and case forms and the second with clause types and verb classes.

The tagmemic model of grammar is based on the concept that syntax is organised in hierarchical levels, and that the basic unit of analysis at each level is the tagmeme, a unit consisting of a slot and filler class. This study concentrates on the tagmemes that constitute the clause level of Lao. Current tagmemic theory, drawing upon the concepts of case grammar, recognises that the grammatical slots of this level have overt syntactic relationships (case forms) and covert semantic relationships (case relations) with the predicative verb, which is considered to be the central slot of the clause (Pike 1970 and Liem 1975). This study will attempt to define these properties for the clause tagmeme of Lao. The occurrence of Fillmore's (1968) set of case relations and case forms

along with additional case forms posited by Liem (1975 and 1976) will be assigned to clause level tagmemes in the form of syntactic features in accordance with the lexicase model of grammar (Starosta 1971a, 1971b and 1973). At the case level of analysis, the possible co-occurrences in a clause tagmeme of members of the set of case relations and the set of case forms established for Lao will first be determined. Then, clause types will be classified in terms of the case features marking their constituent tagmemes.

The data for this study was obtained through fieldwork with five native speakers of Lao. Two informants used are bilingual in English and Lao, one bilingual in Lao and Thai, and the other two basically monolingual with a limited knowledge of English. The elicitation procedure used with the bilingual speakers was to present each informant with an English sentence and request a translation of it into Lao. The procedure used with the other informants involved using an interpreter to present a situation or event to them and requesting a Lao sentence describing the situation or event from the informants. The system of transcription used in this study is that employed in the text, *Lao Basic Course*, published by the Foreign Service Institute (FSI 1970).

## 1. CASE FORMS AND CASE RELATIONS

### 1.1. LAO CASE RELATIONS AND CASE FORMS

Lao has the following twelve universal case relations, as defined by Fillmore (1968) and Liem (1975 and 1976).

- (1) AGENTIVE (AGT) case: the actant that instigates the action.
- (2) OBJECTIVE (OBJ) case: the actant whose role in the clause is identified by the verb itself.
- (3) DATIVE (DAT) case: the animate actant receiving the action or effect of the verb.
- (4) BENEFACTIVE (BEN) case: the actant receiving the benefit of the verbal action.
- (5) COMITATIVE (COM) case: the actant accompanying another actant in the verbal activity or state.
- (6) INSTRUMENT (INS) case: the inanimate object or force causally involved in the verbal action or state.
- (7) LOCATIVE (LOC) case: the actant indicating the spatial location or area of the verbal action or state.
- (8) DIRECTION (DIR) case: the actant indicating the directional orientation of the verbal action or state.
- (9) TIME (TIM) case: the actant indicating the temporal setting of the verbal action or state.
- (10) SOURCE (SRC) case: the actant indicating the spatial or temporal point from which the verbal action began.

- (11) GOAL (GOL) case: the actant indicating the spatial or temporal end point toward which the verbal action is oriented.
- (12) EXTENT (EXT) case: the actant indicating the spatial or temporal extent through which the verbal action has occurred.

These case relations represent universal semantic relations holding between a predicate and its arguments within a clause in accordance with Fillmore's standard case grammar approach (Fillmore 1968 and 1971). Fillmore's principle of limiting the possible occurrences of each case to one instance per clause will be adhered to in this study for the most part. However, as Starosta (1973) has noted, this constraint seems too strong, in that more than one instance of a single case per clause can be allowed providing that the NP's marked for this case are coreferential. Therefore, in this study, coreferential NP's having the same semantic relation to the verb of the clause, will be marked with the same case relation despite their co-occurrence in the same clause.

There are ten case forms, or case realisations, into which the universal case relations are pigeon-holed. These forms are defined either in terms of position of the noun phrase relative to the verb or in terms of its occurrence with certain prepositions. Since the level of analysis is the clause level, these prepositions will not be assigned case frames as in the standard lexicase model. Only the main verb of each clause will host a case frame, which will specify both the required case relations it takes as well as the case forms with which they may co-occur. Accordingly, the entire prepositional phrase tagmeme rather than the case marking preposition will be marked for case form as well as case relation. This restriction of the level of case analysis to the clause level tagmemes has been made only to facilitate the classification and comparison of Lao clause types for pedagogical purposes. Overt case forms for tagmemes of the clause level, then, are marked by the following characteristics in Lao:

- (1) NM (nominative): marked by the position immediately preceding the verb.
- (2) O (objective): marked by the position immediately following the verb.
- (3) D (dative): marked by the preposition *hâj*.
- (4) B (benefactive): marked by the preposition *sămlāp*.
- (5) C (comitative): marked by the preposition *nám*.
- (6) L (locative): marked by the preposition *jūū*.
- (7) I (instrumental): marked by the preposition *dùaj*.

- (8) D1 (directional): marked by the prepositions máa, paj, khým, lón, wàj, ôk, khâam, khâw, taam, hăa, sāj and mýa.  
 (9) Sr (source): marked by the prepositions câak and tēē.  
 (10) G (goal): marked by the prepositions thǎn and hòt.

The set of prepositions functioning as case markers in Lao consists of both non-derived (intrinsic) prepositions and derived prepositions (coverbs) as defined in Kullavanijaya 1974 and Clark 1975. The derived prepositions or 'coverbs' are the dative marker hāj, the comitative marker nām, the locative marker jūū, the source marker câak, as well as the entire set of direction and goal marking prepositions. Each of these derived prepositions is derivationally related to a homophonous and synonymous verb in the language hosting the same case frame. The direction and goal marking prepositions are related to a set of adverbial particles usually occurring after an object NP as part of a transitive compound verb (FSI 1970). However, neither the verbal nor adverbial occurrence of these forms functions as a case marker. Since in the lexibase model, the syntactic category of a lexical item is determined solely by its structural position in a phrase marker, a separate lexical item must be listed in the lexicon for its occurrence in each kind of construction, unless a totally productive derivational rule can be written as a redundancy rule. Homophonous forms of different grammatical categories, but with identical syntactic-semantic features are related by lexical derivation rules. The markers listed above function as case form markers at the clause level only when marked as prepositions, or coverbs, in the lexicon enabling them to occur in construction with a noun phrase in a prepositional phrase. Therefore, the lexical items listed above must co-occur with another verb (the main verb) and immediately precede a noun phrase in a prepositional phrase tagmeme, in order to mark that tagmeme with the appropriate case form. For example, considering the sentences 1-5 below, only in the second and fifth sentences does paj function as a coverb marking a case form.

- (1) lǎaw lēēn paj.  
 [+NM]  
 [+OBJ]  
 'He ran away.'
- (2) lǎaw lēēn paj wīan̄can.  
 [+NM] [+D1]  
 [+OBJ] [+DIR]  
 'He ran to Vientiane.'
- (3) lǎaw paj sỳy pým.  
 [+NM] [+O]  
 [+OBJ] [PURPOSE]  
 'He went to buy books.'

- (4) láaw paj wíangcan.  
 [+NM] [+O]  
 [+OBJ] [+DIR]  
*'He went to Vientiane.'*
- (5) láaw aw pým paj wíangcan.  
 [+NM] [+D1]  
 [+AGT] [+DIR]  
*'He took books to Vientiane.'*
- (6) láaw aw pým paj.  
 [+NM] [+O]  
 [+AGT] [+OBJ]  
*'He took books.'*

In the first and sixth sentence *paj* is an adverbial particle since it is followed by no NP. In sentence #4 it is a main verb since there is no other verb in the sentence. In sentence #3 it is also a main verb, followed by an embedded purpose clause in the +O case form. The other non-derived or intrinsic prepositions, are termed as such since they are derivationally related to no other forms marked for another grammatical category. That is, they occur only in prepositional slots as *paj* does in sentences #2 and #5. Both types of prepositions, however, function identically as case markers with respect to the main verb of the clause.

Most studies using the lexicase model (Kullavanijaya 1974, Clark 1975 and Manley 1971) have combined the locative, source, goal, and directional case relations into a single case termed 'locative'. Fillmore (1968) determined that it was unnecessary to posit separate cases for these since their distinctions are superficial and can be accounted for by the character of the verb hosting them. Kullavanijaya (1974) and Clark (1975) characterise their differences in terms of inherent syntactic-semantic features marking locative verbs and locative case markers, classifying the latter into markers of sub-case forms. The features used in their analysis are: [ $\pm$ dir] - (direction of motion or action), [ $\pm$ gol] - (extrapolated end point of directional action), [ $\pm$ src] - (starting point of directional action), and [ $\pm$ terminus] - (actual end point of +gol action). Lexical redundancy rules establish the relationship among these features as reflected in their definitions above. In this study, however, the semantic functions grouped together in the locative case relation in the analyses mentioned will be distributed among the LOC, DIR, GOL, SRC, and EXT case relations as previously defined in order to simplify the semantic distinctions determined by these features for contrastive analysis. Likewise, the sub-case forms established by Kullavanijaya's and Clark's semantic-syntactic features will be maintained as separate case forms to more clearly

illustrate the semantic functions of their case markers. The forms distinguished here, nevertheless, can be clearly correlated with the sub-case forms in terms of features: +L - [-dir], +D [+dir, +gol, (-ter)], +SR - [+dir, +src], and +G - [+dir, +ter].

The definitions of the cases +SRC, +GOL, and +EXT indicate that either temporal or spatial relationships can be represented by these cases. The definition of +TIM states that a temporal relationship is inherent in this case, whereas the +LOC and +DIR cases indicate inherently spatial relationships. An examination of the possible co-occurrences of case forms and case relationships reveals that although a case form is listed as realising a particular case relation, when there is more than one case marker for that form, only a specified subset of these markers can realise a temporal or a spatial relation. These restrictions indicate the need to establish sub-case forms classified by the syntactic-semantic features [+temporal] and [+spatial]. Case markers will be assigned these features according to their ability to mark forms realising temporal or spatial sub-case relations. The case relations +LOC, +DIR, +SRC, +GOL (since +TIM and +EXT occur only in the +O form in Lao, they will not be considered here, although this may not be true for other languages) co-occur with +L, +D, +SR, and +G markers which must be assigned these syntactic-semantic selectional features as a restriction on their co-occurrence with NP's with these semantic features.

jūū	[-tem, +spa]	taam	[-tem, +spa]
paj	[-tem, +spa]	hǎa	[+tem, +spa]
mǎa	[-tem, +spa]	sāj	[-tem, +spa]
khŷn	[-tem, +spa]	myá	[-tem, +spa]
lón	[-tem, +spa]	câak	[+tem, +spa]
wàj	[-tem, +spa]	tēē	[+tem, +spa]
ôk	[-tem, +spa]	thǎn	[-tem, +spa]
khâam	[-tem, +spa]	hòt	[+tem, +spa]
khâw	[-tem, +spa]		

In Lao, it seems that the more general usage of prepositions, with both temporal and spatial relations, is restricted to a small set, basically +G and +SRC markers, with the exception of hǎa. It is probable that the +tem usage came about due to the metaphorical extension of the basic [+spa] usage to a limited set of case form markers in accordance with the localist theory (Anderson 1971). However, this question is far beyond the scope of this paper. The features proposed will be used only to illustrate this restriction on case marker usage. Although this restriction occurs at the phrase level, it has been briefly discussed because of its pedagogical usefulness.

There are 22 possible combinations of case relations and case forms in Lao as illustrated in the two-dimensional matrix chart below:

	NM	O	D	B	C	L	I	D1	Sr	G
AGT	1									
OBJ	2	5								
DAT	3		12							
BEN			13	14						
COM					15					
INS	4	6			16		17			
LOC		7				19				
DIR		8						20		
TIM		9								
SRC							18		21	
GOL		10								22
EXT		11								

These possible combinations of case forms and case relations can be represented in terms of features marking noun phrases that fill clause level tagmemes in accordance with the lexicase model (Starosta 1973). The following example sentences illustrate the co-occurrence of case forms and relations in Lao, as listed in the above chart.

- (1) [+NM, +AGT] is hosted by transitive, agentive verbs.

(1.1) láaw sỳy nangsỳyphím.  
 [+NM] [+O]  
 [+AGT] [+OBJ]  
 'He bought a newspaper.'

- (2) [+NM, +OBJ] is hosted by copulative, stative, and intransitive verbs, and the submissive verb thyyk.

(2.1) láaw mēēn najmᵃᵃ.  
 [+NM] [+O]  
 [+OBJ] [+OBJ]  
 'He is a doctor.'

(2.2) láaw sàa lǎaj.  
 [+NM]  
 [+OBJ]  
 'He is very slow.'

- (2.3) láaw naaw láaj.  
 [+NM]  
 [+OBJ]  
*'He is very cold.'*
- (2.4) láaw taaj lèew.  
 [+NM]  
 [+OBJ]  
*'He is dead already.'*
- (2.5) láaw jūw wíangcan.  
 [+NM] [+O]  
 [+OBJ] [+LOC]  
*'He lives in Vientiane.'*
- (2.6) láaw paj wíangcan.  
 [+NM] [+O]  
 [+OBJ] [+DIR]  
*'He went to Vientiane.'*
- (2.7) láaw máa hòot wíangcan.  
 [+NM] [+G1]  
 [+OBJ] [+GOL]  
*'He arrived in Vientiane.'*
- (2.8a) láaw thâyk paj.  
 [+NM] [+O]  
 [+OBJ] [+INS]  
*'He was forced to go.'*
- (2.8b) láaw thâyk khaw dii.  
 [+NM] [+O]  
 [+OBJ] [+INS]  
*'He was beaten by them.'*
- (2.8c) lōt láaw thâyk fáj māj.  
 [+NM] [+O]  
 [+OBJ] [+INS]  
*'His car caught fire (burned).'*
- (2.8d) lōt láaw thâyk khaw lāk lèew.  
 [+NM] [+O]  
 [+OBJ] [+INS]  
*'They stole his car.'*
- (3) [+NM, +DAT] is hosted by transitive dative verbs.
- (3.1a) láaw huàcák láaw.  
 [+NM] [+O]  
 [+DAT] [+OBJ]  
*'He knows her.'*
- (3.1b) láaw hòu láaw paj lèew.  
 [+NM] [+O]  
 [+DAT] [+OBJ]  
*'He knows that she went already.'*



- (3.1c) láaw jâak hâj láaw paj lèew.  
 [+NM] [+O]  
 [+DAT] [+OBJ]  
*'He wanted her to go.'*

Liem (1975) maintains that the combination [+NM, +LOC] occurs in the following clause type:

- hông nii naaw lăaj.  
 [+NM]  
 [+LOC]  
*'This room is very cold.'*

However, most studies in case grammar theory and analysis maintain that the [+OBJ] relation must be present in every clause. Therefore the features [+NM, +OBJ] will be assigned to the noun phrase above. The semantic distinction of this sentence is best marked at the word level with the syntactic-semantic feature [+location] marking the word hông.

- (4) [+NM, +INS] is hosted by transitive, agentive verbs, and by intransitive verbs with 'potent' nouns.

- (4.1) mît duan nii tát sîn.  
 [+NM] [+O]  
 [+INS] [+OBJ]  
*'This knife cuts meat.'*

- (4.2) lôt nii paj wíŋcan. (Only 'potent' nouns can occur in  
 [+NM] [+O] this slot.)  
 [+INS] [+OBJ]  
*'This car went to Vientiane.'*

- (5) [+O, +OBJ] is hosted by transitive verbs.

- (5.1) láaw sỳy pỳm.  
 [+NM] [+O]  
 [+AGT] [+OBJ]  
*'He bought books.'*

- (5.2) láaw hùu lỳan nân.  
 [+NM] [+O]  
 [+DAT] [+OBJ]  
*'He knows that story.'*

- (6) [+O, +INS] is hosted only by certain transitive verbs and by the submissive verb thyyk.

- (6.1) láaw sàj mît lăaj.  
 [+NM] [+O]  
 [+AGT] [+INS]  
*'He used knives a lot.'*

- (6.2) láaw thyyk khaw paj.  
 [+NM] [+O]  
 [+OBJ] [+INS]  
*'He was forced to go.'*

- (7) [+0, +LOC] is hosted by intransitive, locative verbs.

(7.1) láaw jūw wíʔanʔan.  
 [+NM] [+O]  
 [+OBJ] [+OBJ]  
*'He lives in Vientiane.'*

- (8) [+0, +DIR] is hosted by intransitive, directional verbs.

(8.1) láaw paj wíʔanʔan.  
 [+NM] [+O]  
 [+OBJ] [+DIR]  
*'He went to Vientiane.'*

- (9) [+0, +TIM] can be hosted by any verb, since it is a sentence level tagmeme.

(9.1) láaw paj mỳy wāān nii.  
 [+NM] [+O]  
 [+OBJ] [+YIM]  
*'He went yesterday.'*

- (10) [+0, +GOL] is hosted by intransitive +GOL verbs.

(10.1) láaw máa wíʔanʔan.  
 [+NM] [+O]  
 [+OBJ] [+GOL]  
*'He arrived in Vientiane.'*

- (11) [+0, +EXT] is hosted by any verb not marked otherwise.

(11.1) láaw paj sǎʔŋ mỳy.  
 [+NM] [+O]  
 [+OBJ] [+EXT]  
*'He went for two days.'*

(11.2) láaw paj sǎʔŋ lak.  
 [+NM] [+O]  
 [+OBJ] [+EXT]  
*'He went for two miles.'*

- (12) [+D, +DAT] is hosted by transitive agentive verbs.

(12.1) láaw khâaj pỳm hâj khôj.  
 [+NM] [+O] [+D]  
 [+AGT] [+OBJ] [+DAT]  
*'He sold books to me.'*

- (13) [+D, +BEN] is hosted by transitive and intransitive verbs and is structurally ambiguous with sentences containing [+D, +DAT],

(13.1) láaw sỳy pỳm hâj khôj.  
 [+NM] [+O] [+D]  
 [+AGT] [+OBJ] [+BEN]  
*'He bought books for me.'*

- (13.2) láaw paj hâj khôj.  
           [+NM]                   [+D]  
           [+OBJ]               [+BEN]  
           *'He went for me.'*

- (14) [+B, +BEN] is hosted only by the verb mfi wâj 'to be used'. The meaning associated with the use of this case form is 'purpose' or 'for the use of' rather than 'for the benefit of' as in the sentences above. Consequently, the filler class of this tagmeme consists of nouns derived from verbs.

- (14.1) pyyn mfi wâj samlāp njín.  
           [+NM]                   [+B]  
           [+OBJ]               [+BEN]  
           *'Guns are for shooting.'*

- (15) [+C, +COM] can occur with any verb.

- (15.1) láaw paj nám khôj.  
           [+NM]                   [+C]  
           [+OBJ]               [+COM]  
           *'He went with me.'*

- (16) [+C, +INS] is hosted by intransitive verbs.

- (16.1) láaw paj nám lōt.  
           [+NM]                   [+C]  
           [+OBJ]               [+INS]  
           *'He went by bus.'*

- (17) [+I, +INS] is hosted by transitive and intransitive verbs.

- (17.1) láaw sỳy pỳm dùaj laáw ɲən.  
           [+NM]           [+O]                   [+I]  
           [+AGT]           [+OBJ]               [+INS]  
           *'He bought books with Lao money.'*

- (17.2) láaw dæntháaŋ dùaj khỹaŋ bin.  
           [+NM]                   [+I]  
           [+OBJ]               [+INS]  
           *'He travels with (by means of) aeroplanes.'*

- (18) [+I, +SRC] is hosted by the stative verb hēt 'to be made'.

- (18.1) p55ŋjiam n̄i hēt dùaj kèɛw.  
           [+NM]                   [+I]  
           [+OBJ]               [+SRC]  
           *'This window is made of glass.'*

- (19) [+L, +LOC] is hosted by any verb.

- (19.1) láaw sỳy pỳm jūū amerika.  
           [+NM]           [+O]                   [+L]  
           [+AGT]           [+OBJ]               [+LOC]  
           *'He bought books in America.'*

- (20) [+D1, +DIR] is hosted by transitive verbs and by intransitive verbs when the [+D] marking prepositions or coverbs occur in a clause.

(20.1) láaw      aw pým      paj wíañcan.  
           [+NM]        [+O]        [+D1]  
           [+AGT]        [+OBJ]        [+DIR]  
           *'He brought books to Vientiane.'*

(20.2) láaw      njāān paj wíañcan.  
           [+NM]                    [+D1]  
           [+OBJ]                    [+DIR]  
           *'He walked to Vientiane.'*

- (21) [+Sr, +SRC] is hosted by all verbs if it is [+tem], and by most verbs if [+spa].

(21.1) láaw      máa cāak amerika.  
           [+NM]                    [+Sr]  
           [+OBJ]                    [+SRC]  
                                       [-tem]  
           *'He came from America.'*

(21.2) láaw      paj lèew tēē mỳy      wāan nīi.  
           [+NM]                    [+Sr]  
           [+OBJ]                    [+SRC]  
                                       [+tem]  
           *'He has been gone since yesterday.'*

- (22) [+G, +GOL] is hosted by all verbs if [+tem] and by most if [-tem].

(22.1) láaw      paj thəŋ wíañcan.  
           [+NM]                    [+G]  
           [+OBJ]                    [+GOL]  
                                       [-tem]  
           *'He went to (and arrived at) Vientiane.'*

There is no +E case form in Lao as in Vietnamese. The case relation +EXT is realised in the +O case form only.

- (23) In addition to the case relations discussed, adjectival phrase tagmemes denoting 'manner' function like nominal case relations in respect to a clause level predicate, and seem to take case relations in the same fashion as the cases above. The [+MANNER] slot is filled with an adjective (or stative verb) phrase realised as the +O case form.

(23.1) láaw      paj sàa.  
           [+NM]                    [+O]  
           [+OBJ]                    [+MANNER]  
           *'He went slowly.'*

Liem (1975) assigns the case-like relation [+PURPOSE] to embedded sentences forming surface purpose phrases in his analysis of Vietnamese. However, as example #14 illustrates, purpose can be conveyed in Lao by a tagmeme marked [+B, +BEN] filled with the preposition sāmīāp and a derived noun. To be consistent in this analysis, purpose phrase tagmemes

will be marked [+BEN] when realised by the +O case form relative to the higher verb, as well. This eliminates the need to set up a separate case-like relation to account for these phrases.

lǎaw paj sỳy nansỳyphím.  
 [+NM] [+O]  
 [+AGT] [+BEN]  
 'He went to buy newspapers.'

1.2. The case realisations of Lao seem to have the following relative order:

			+I		+G	+L
+NM	+Verb	+O	+D	+Sr	+Di	+C

Although it is theoretically possible for all of the case forms to occur in a single clause, it does not seem likely that more than four (or five if one is sentence level rather than clause level) ever do in actual speech. Those sentences containing more than three case marked tagmemes were considered awkward by the informants and tended to be rejected in favour of sentences consisting of two or more clauses.

The occurrence of case forms is restricted to one instance per clause, generally, for either syntactic or performance reasons. There is a syntactic restriction against the co-occurrence of more than one +NM, +I, and +B slot. When two case relations are neutralised by a single case form, two occurrences of that form are allowed if they realise different relations. This occurs in Lao with the +C and +D cases.

lǎaw paj nám lōt nám khôj.  
 [+NM] [+C] [+C]  
 [+OBJ] [+INS] [+COM]  
 'He went by bus with me.'

lǎaw khâaw pỳm hāj hākhían hāj khôj.  
 [+NM] [+O] [+D] [+D]  
 [+AGT] [+OBJ] [+DAT] [+BEN]  
 'He sold the books to the students for me.'

In both cases, however, there seem to be order restrictions requiring the case relation Fillmore classifies as 'inner' to occur before the 'outer' case relation (Fillmore 1971). In other words, the +INS and +DAT must precede the +COM and +BEN cases respectively when they take the same form within a single clause.

Also more than one 'locative' case form (+L, +Di, +Sr, +G) providing they are coreferential (inclusive) (Starosta 1971). When there is more than one occurrence of these cases, the included occurrence must precede the occurrence with more general reference.

lǎaw sǎon jūū hǒonghǎn jūū wíangcan.  
 [+NM] [+L] [+L]  
 [+OBJ] [+LOC] [+LOC]

'He teaches at a school in Vientiane.'

When two +O case forms occur in a single clause, the second occurrence must be either the +BEN purpose clause or the +MAN adverbial.

lǎaw khǎaw pým waj.  
 [+NM] [+O] [+O]  
 [+AGT] [+OBJ] [+MAN]

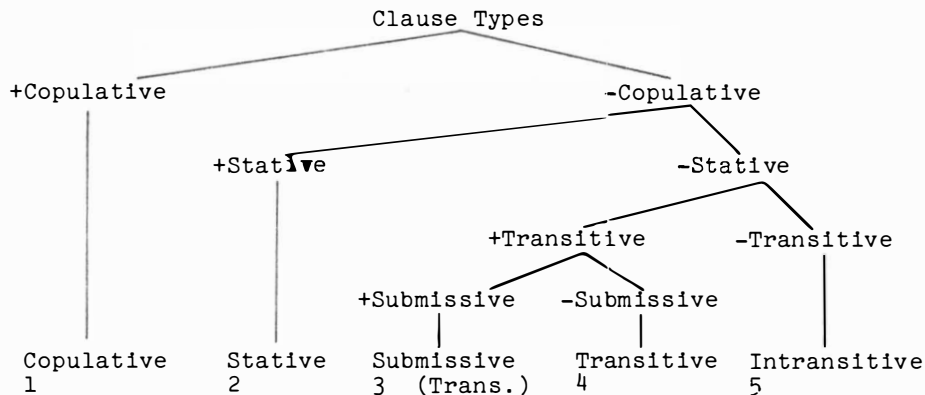
'He sells books quickly.'

lǎaw khǎaw pým sýy kin.  
 [+NM] [+O] [+O]  
 [+AGT] [+OBJ] [+BEN]

'He sells books to buy food.'

## 2. CLAUSE TYPES AND VERB CLASSIFICATION

Lao clauses contain at least an obligatory nucleus predicate tagmeme and an obligatory nucleus nominative tagmeme. There are five types as classified in the chart below.



### 2.1. CLAUSE TYPES

#### 2.1.1. Copulative Clause

[[ -NM, +OBJ ] [+Verb, +Copula] [+O, +OBJ]]

#### 2.1.2. Stative Clause

[[ +NM, +OBJ ] [+Verb, +Stative]]

#### 2.1.3. Submissive Clause

[[ +NM, +OBJ ] [+Verb, +Trans., +Submiss.] [+O, +INS]]

The [+O, +INS] in this clause type must be filled by a clause. Only one verb, thÿyk, can fill the +Verb slot.

#### 2.1.4. Transitive Clause

[ [+NM, +DAT]  
[ [+NM, +AGT] [+Verb, +Trans.] [+O, +OBJ]]  
[ [+NM, +INS]

#### 2.1.5. Intransitive Clause

[ [+NM, +OBJ] [+Verb, +Trans.]]

### 2.2. CLASSIFICATION OF VERBS

The verbs of Lao can be classified into types according to the case frames they have. The case frames occurring with each group along with a few sample verbs from that group are given below.

#### 2.2.1. mĕĕn 'be'

This class consists of copula verbs. It hosts the [+NM, +OBJ] and [+O, +OBJ] tagmemic slots:

[ [+NM, OBJ] _____	[ +O, +OBJ]]
noun phrase	noun phrase
verb phrase	verb phrase
clause	clause
prep. phrase	prep. phrase

The [+NM, +OBJ] and [+O, +OBJ] slots must be filled by the same grammatical category, and if the category is a noun phrase, the fillers must be coreferential:

lǎaw	mĕĕn	najmǎw
[+NM]		[+O]
[+OBJ]		[+OBJ]
[+NP]		[+NP]

*'He is a doctor.'*

paj	mĕĕn	tǎw	taj.
[+VP]		[+VP]	

*'To go is to die.'*

lǎaw	paj	mĕĕn	kwəŋ	salaa	kǎw	lĕew.
[+clause]				[+clause]		

*'He was intelligent to go.'*

câak	mí	paj	tyyŋ	hân	mĕĕn	sǎw	lak.
[+PP]						[+PP]	

*'From here to there is two miles.'*

The only other verb in this class is pen which must occur with the +NM and +O slots filled with noun phrases marked +human.

## 2.2.2. sàa 'slow'

This class of verbs is stative and can be adverbialised. Its case forms contain a [+NM, +OBJ] slot and it can take an intensifier.

láaw sàa lǎaj.  
 [+NM] [+intens.]  
 [+OBJ]  
*'He is very slow.'*

láaw paj saa laaj.  
 [+NM] [+O] [+intens.]  
 [+OBJ] [+MAN]  
*'He went very slowly.'*

Other verbs in this class include waj 'fast', sua 'bad', and dii 'good'.

## 2.2.3. naaw 'cold'

The verbs in this class occur in a case frame containing [+NM, +OBJ]. They are stative and can take an intensifier, but cannot be adverbialised.

láaw nǎaw laaj.  
 [+NM] [+intens.]  
 [+OBJ]  
*'He is very cold.'*

Other verbs in this class include diicaj 'happy', sǎaw 'sad', and khâj 'sick'.

## 2.2.4. taaj 'to be dead'

This class consists of stative verbs, hosting a [+NM, +OBJ] slot, and which cannot take an intensifier.

láaw taaj.  
 [+NM]  
 [+OBJ]  
*'He is dead.'*

Also in this class are the verbs sivit 'to be alive', and kêet 'to be born'.

## 2.2.5. thâyk 'to be affected'

There is only one verb in this class, which hosts [+NM, +OBJ] and [+O, +INS] tagmemes. The latter slot must be filled with a clause.

[+NM, +OBJ] ——— [+O, +INS]  
 [+NP] [+clause]



lǎaw	thỳk	paj.
[+NM]		[+O]
[+OBJ]		[+INS]
[+NP]		[+clause]

*'He was forced to go.'*

lōt	thỳk	kāk	lèew.
[+NM]		[+O]	
[+OBJ]		[+INS]	
[+NP]		[+clause]	

*'The car was already stolen.'*

In the lesser urbanised dialects of Lao, this verb can host only an animate +NM noun like the Vietnamese submissive verbs *bi* and *được*. However, due to the influence of Western languages, particularly French, the usage of *thỳk* has been extended to non-animate subjects to parallel the passive constructions found in European languages. Therefore, in Lao, the case frame for *thỳk* contains a [+NM, +OBJ] slot rather than a [+NM, +DAT] slot as the Vietnamese submissive verbs *bi* and *được*.

#### 2.2.6. sỳy 'buy'

This class of ditransitive benefactive verbs has been the case frame [+NM, +AGT], [+O, +OBJ], and [+D, +BEN].

lǎaw	sỳy	pỳm	hāj	khôj.
[+NM]		[+O]		[+D]
[+AGT]		[+OBJ]		[+BEN]

*'He bought books for me.'*

#### 2.2.7. khâaw 'sell'

This class of verbs is characterised by the case frame consisting of the tagmemes [+NM, +AGT], [+O, +OBJ], and [+D, +DAT].

lǎaw	khâaw	pỳm	hāj	khôj.
[+NM]		[+O]		[+D]
[+AGT]		[+OBJ]		[+DAT]

*'He sold books to me.'*

Due to the neutralisation of the +DAT and +BEN relations in the +D case form, the above sentence is ambiguous. The phrase *hāj khôj* can also be marked [+D, +BEN] in a sentence with a deleted [+D, +DAT] phrase.

#### 2.2.8. lỳak 'choose'

This class of transitive verbs is characterised by the case frame [+NM, +AGT] and [+O, +OBJ], the latter of which can be complementised by an embedded copulative sentence.



láaw	hùu	láaw	paj	lèew.
[+NM]		[+O]		
[+DAT]		[+OBJ]		
		[+clause]		

'He knows that she already went.'

láaw	hùu	láaw.
[+NM]		[+O]
[+DAT]		[+OBJ]

'He knows her.'

Other verbs in this class are dàj njín 'hear', kâwcaj 'understand', and khĩt 'think'.

2.2.12. sán 'hate'

These verbs take the same case frame as *hùu*, but differ in that they can be modified by an intensifier. The [+0, +OBJ] tagmeme can be a dependent clause.

[[+NM, +DAT] (intensifier) ——— [ +O, +OBJ]]  
noun phrase  
dep. clause

l <sup>á</sup> aw	l <sup>ǎ</sup> aj	sán	l <sup>á</sup> aw.
[+NM]	[+intens]		[+O]
[+DAT]			[+OBJ]

'He hates her a lot.'

Other verbs in this class include *hāk* 'love', *iiton* 'pity', and *aa* 'shy, ashamed'.

2.2.13. jūū 'live at, be at'

The locative verb, *juu*, is the only member of this class. It can serve as a main verb or as a coverb marking the +L, locative, case form. As a main verb it hosts [+NM, +OBJ] and [+O, +LOC] tagmemes.

l áaw	j ūū	w íaŋcan.
[+NM]		[+O]
[+OBJ]		[+LOC]

'He is in Vientiane'.

## 2.2.14. ȡȡk 'go out'

The verbs in this class are intransitive and are characterised by being directional and having the [+locomotion] feature. They host [+NM, +OBJ] and [+O, +DIR] tagmemes.

l áaw	ôok	hôoη.
[+NM]		[+O]
[+OBJ]		[+DIR]

'He went out of the room.'

These verbs are derivationally related to a set of homophonous co-verbs marking the directional, +D1, case form. Other verbs in this class include *lống* 'go down', *khộn* 'go up', and *khâm* 'go through'.

#### 2.2.15. *paj* 'go'

Verbs in this class host the [+NM, +OBJ] and [+O, +DIR] case frame. They are distinguished as [-locomotion].

*lấaw paj wíangcan.*  
 [+NM] [+O]  
 [+OBJ] [+DIR]  
*'He went to Vientiane.'*

These verbs are derivationally related to a set of coverbs marking the +D1 case form. The other verb in this class is *máa* 'come'.

#### 2.2.16. *thắg* 'arrive'

This class of verbs is marked by the feature [+goal], and the case frame [+NM, +OBJ] and [+O, +GOL]. The verbs in this class are derivationally related to a set of coverbs marking the +G case form.

*lấaw thắg wíangcan.*  
 [+NM] [+O]  
 [+OBJ] [+GOL]  
*'He arrived in Vientiane.'*

2.2.17. There does not seem to be any verb class functioning as an existential sentence introducer similar to the Vietnamese verb *có*.

BIBLIOGRAPHY

ANDERSON, John M.

- 1971 *The Grammar of Case: Towards a Localistic Theory*. London: Cambridge University Press.

CLARK, Marybeth

- 1975 Coverbs and Case in Vietnamese. Doctoral dissertation, University of Hawaii, Honolulu. Published 1978 as PL, B-48.

FILLMORE, Charles J.

- 1968 'The Case for Case'. In: E. Bach and R.T. Harms, eds *Universals in Linguistic Theory*, 1-88. New York: Holt, Rinehart and Winston.
- 1971 'Some Problems in Case Grammar'. *Ohio State Working Papers in Linguistics* 10:245-65.

KULLAVANIJAYA, Pranee

- 1974 Transitive Verbs in Thai. Doctoral dissertation, University of Hawaii.

LIEM, Nguyen Dang

- 1971a 'Clause Units in English and Southeast Asian Languages in Contrast'. In: Gerhard Nickel, ed. *Papers in Contrastive Linguistics*, 157-74. Papers from the Second International Conference of Applied Linguistics, Stuttgart. London: Cambridge University Press.
- 1971b 'Modern Linguistic Theories and Contrastive Linguistics'. In: Gerhard Nickel and Albert Raasch, eds *Kongressbericht*

*der 3 Jahrestagung der Gesellschaft für angewandte Linguistik.* Heidelberg: Julius Groos.

- 1974 Clause Units in South-East Asian Languages (Burmese, Cambodian, Cantonese, Lao, Thai, and Vietnamese)'. In: Luigi Heilmann, ed. *Proceedings of the Eleventh International Congress of Linguists, Bologna Florence, Aug.28-Sept.2, 1972*, 231-52. Bologna: Mulino.
- 1975 *Cases, Clauses and Sentences in Vietnamese.* PL, B-37.
- 1976 'Cases and Clauses in Vietnamese'. In: Philip N. Jenner, Lawrence C. Thompson and Stanley Starosta, eds. *Austronesian Studies. Oceanic Linguistics Special Publication No.13*: 773-99. 2 vols. Honolulu: University Press of Hawaii.

NGUYEN DANG LIEM *see* LIEM, Nguyen Dang

PLATT, Heidi K.

- 1970 A Comparative Study of English and German. Ph.D. dissertation, Monash University, Clayton, Victoria.

PLATT, John T.

- 1971 *Grammatical Form and Grammatical Meaning.* Amsterdam: North-Holland.

STAROSTA, Stanley

- 1971a 'Derivation and Case in Sora Verbs'. *Indian Linguistics* 32/3:194-206.
- 1971b 'Lexical Derivation in a Case Grammar'. *Working Papers in Linguistics, University of Hawaii* 3/8:83-101.
- 1973 'The Faces of Case'. *Language Sciences* 25:1-14.

## A CLASSIFICATION OF THAI CLASSIFIERS

SamAng HIRANBURANA

One area which has hardly been touched and which has always baffled a non-Thai speaker is the use of classifiers. At first impression there seems to be one classifier for each noun. My counting, which I believe is fairly extensive, reveals 75 in all, not counting those classifiers used as units in measurement of weight, distance and time, nor those referred to by U. Warotamasikkhadit (1963) as 'reduplicative classifiers', which is the use of the noun itself as its own classifier such as /my<sup>1</sup> sɔɔ<sup>5</sup> my<sup>1</sup>/ 'two hands',<sup>1</sup> nor classifiers for the non-concrete nouns such as /wæp<sup>2</sup>/ 'a flash', /kæp<sup>2</sup>/ 'a rap'.

The fullest account of the use of Thai classifiers can be found in Noss 1964. It is a word which occurs with a noun when a number and/or a demonstrative is also present - /buri<sup>1</sup> sɔɔ<sup>5</sup> muh<sup>1</sup>/ 'two cigarettes' or /buri<sup>1</sup> muh<sup>1</sup> ni<sup>1</sup>/ 'this cigarette'. However, in the context where the noun is understood, the noun may be omitted, leaving its classifier to act as the head of a classifier phrase - /sɔɔ<sup>1</sup> si<sup>5</sup> lya<sup>5</sup>/ 'the yellow pack'. Noss recognises that there are well over 200 classifiers in Thai. He groups them into four types:

1. A unit classifier which has a specific relationship with one or more concrete nouns.
2. A metric classifier which occurs in enumeration. It has no special relationship with the noun in the way that unit classifiers do but it is used for measurement of a mass noun into fixed unit or conventional contents of a container such as in /khaa<sup>3</sup> saam<sup>5</sup> tcah<sup>1</sup>/ 'three platefuls of rice'. Noss also divides the metric classifiers into five subclasses namely (1) distance and size, (2) weight, (3) container, (4) value, and (5) time.

3. A general classifier which occurs in enumerations after an extremely wide range of nouns with no special relationships with those nouns. Some of them can occur even after abstract nouns. The class is small and probably closed. Noss notes that 14 classifiers belong to this type.

4. An imitative classifier which is the same type of classifiers referred to by Udom Waritamasikkhadit (1963) as 'reduplicative classifiers' which has been mentioned previously.

In his classification, Noss is primarily concerned with the function and usage of the classifiers. There is little attempt to relate classifiers semantically with the nouns they represent. In Noss' opinion (1964), "a far better lexicographical technique, used by Mary Haas and others, is to give the unit classifier in parentheses after each concrete noun listed" and that "The semantic connection between classifier and noun may be buried so deep in history that it makes no sense descriptively".

Obviously, Noss' despair over Thai classifiers can only leave the Thai-learner at a loss, with his memory the only tool available to tackle this area of Thai language. It is my contention, however, that there is a definite semantic connection between a noun and its classifier, and that on being confronted with an unknown object and given its function, most Thais will more or less agree as to what classifier they would use if one is required.

It seems that one may group Thai 'specific classifiers' into three major classes with their subclasses. These, with specifications by various lexical features will give selectional rules which will match with the feature specifications of a noun and permit the noun to select its appropriate classifier. The three major classes are as follows:

1. Classifiers<sub>(amount)</sub>
  - 1.1. Classifiers<sub>(partitive)</sub>
  - 1.2. Classifiers<sub>(pair)</sub>
  - 1.3. Classifiers<sub>(plural)</sub>
2. Classifiers<sub>(shape)</sub>
  - 2.1. Linear
  - 2.2. Quadrilateral
  - 2.3. Circular
  - 2.4. Cylindrical

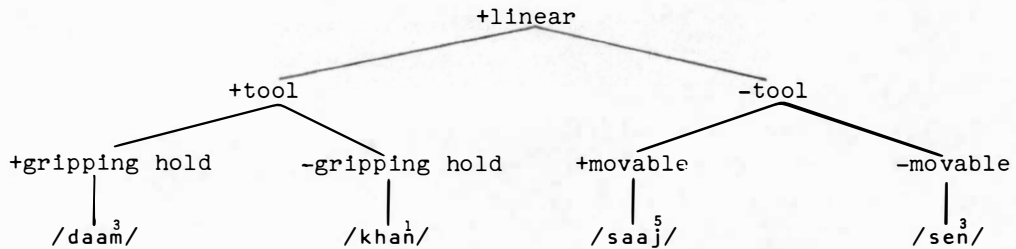


3. Classifiers<sub>(function)</sub>

- 3.1. Weapon
- 3.2. Container
- 3.3. Vehicle
- 3.4. Machinery
- 3.5. Inhabitation
- 3.6. Tool

In this paper, I will use only one area to illustrate my contention namely the use of Classifiers<sub>(shape)</sub> in which there is a definite connection between the classifier and the shape of the noun it is to go with.

## 2.1. Linear Classifiers



Nouns used with each classifier:<sup>2</sup>

- /leṃ³/: /miit³/ (and its noun compounds), /tcɔɔb²/, /siaṃ⁵/, /khem⁵/, /kankraj¹/, /thian¹/, /paǵ¹/ ('oar'), /paǵ¹/ ('spatula'), /khwaan⁵/, /khraad³/, /taliṃ⁵/
- /daam³/: /pkat⁴/, /paakkaṃ⁴/, /diṃ²/
- /khañ¹/: /thanuu⁵/, /naamaǵ⁴/, /rom³/, /bet²/, /tchɔɔñ⁴/, /sɔɔṃ³/, /thappi¹/, /sɔɔ¹/
- /saaǵ⁵/: /thanoṃ⁵/, /thaañ¹/, /mænaam⁴/, /lamthaan¹/, /khlɔɔñ¹/
- /señ³/: /saaǵ⁵ phaan¹/, /bami²/, /wuñ señ³/, /maj⁵/, /faaǵ³/, /khem⁵ khaṭ²/, /tchyaṅ² phuuk rɔɔñ¹ thaaw⁴/, /phom⁵/, /luaṭ³/, /daaǵ³/, /tcyak³/, /leḱ²/ ('steel wire'), /saaǵ⁵ thoorasap²/ ('telephone wire'), /saaǵ⁵ faj¹/ ('electrical wire')

Before discussion of the classifiers and their corresponding nouns, one should mention something about the features used here. The feature [+linear] is used here to imply that length is the striking feature in

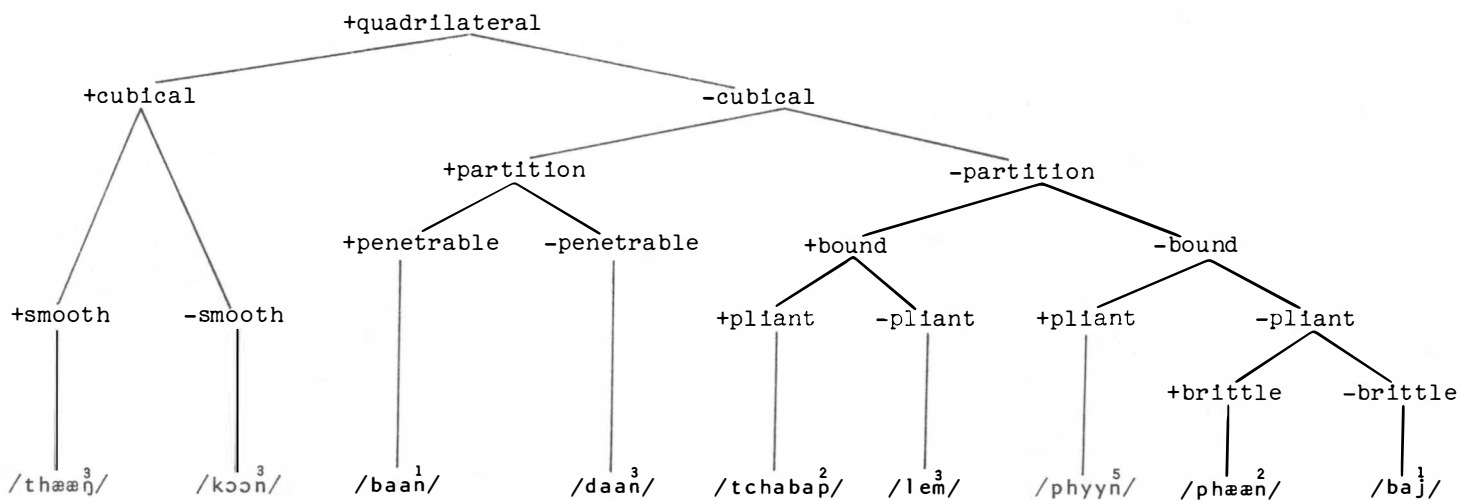
the object which in some cases, such as /thanon<sup>5</sup>/ 'road', can have width as well.

/lem<sup>3</sup>/ is the general classifier used for a linear tool which has to be used with a gripping hold such as /mii<sup>3</sup>/ 'knife', /khem<sup>5</sup>/ 'needle'. /thian<sup>1</sup>/ 'candle' is perhaps the only one exception since when one holds a candle, it need not be a grip. Three of the nouns used with /lem<sup>3</sup>/, namely /pha<sup>4</sup>/ 'fan', /paakkaa<sup>1</sup>/ 'pen' and /diw<sup>3</sup>/ 'a kind of Chinese weapon' can be used alternatively with another classifier /daam<sup>3</sup>/.

When the tool is used without the gripping hold, /khan<sup>1</sup>/ is used. When one holds those nouns, the hold itself is more like a support (e.g. /rom<sup>3</sup>/ 'umbrella') or a counterweight effort (e.g. /thanu<sup>1</sup>/ 'bow') rather than a grip. It is quite certain that this classifier is derived from the same root as two other words: /khaan<sup>1</sup>/ 'to support' and /khaan<sup>4</sup>/ 'to go against'.

/saa<sup>5</sup>/ and /sen<sup>3</sup>/ are non-tool linear classifiers. The basic difference between them is marked by the features [+stationary] and [-stationary]. It may be argued that /saa<sup>5</sup>/ is also used for /thor<sup>1</sup> sa<sup>2</sup>/ 'telephone' and /fa<sup>1</sup> faa<sup>4</sup>/ 'electricity', but this usage in fact refers to the telephone and electricity connections and not to the wire which will be used with /sen<sup>3</sup>/ if a classifier is required.

## 2.2. Quadrilateral Classifiers



Nouns used with each classifier:

- /thæŋ<sup>3</sup>/<sub>1</sub>: /lek<sup>2</sup>/, /thoŋ<sup>1</sup>/, /jaaŋ<sup>1</sup> loŋ<sup>4</sup>/
- /kooŋ<sup>3</sup>/: /hiŋ<sup>5</sup>/, /diŋ<sup>1</sup>/, /khi<sup>3</sup>/, /meek<sup>3</sup>/, /khanom<sup>5</sup> paŋ<sup>1</sup>/, /ʔit<sup>2</sup>/, /naam<sup>4</sup> khæŋ<sup>5</sup>/, /sabu<sup>2</sup>/, /naam<sup>4</sup> taan<sup>1</sup>/, /diŋ<sup>1</sup> niaw<sup>5</sup>/
- /baan<sup>1</sup>/: /pratu<sup>1</sup>/, /naa<sup>3</sup> taan<sup>2</sup>/, /katco<sup>2</sup>ŋaaw<sup>1</sup>/
- /daan<sup>3</sup>/: /kamphæŋ<sup>1</sup>/, /rua<sup>4</sup>/, /faa<sup>5</sup>/, /daan<sup>3</sup>/
- /tchabap<sup>2</sup>/: /tco<sup>2</sup> maa<sup>5</sup>j/, /nan<sup>5</sup>sy<sup>5</sup> phi<sup>1</sup>m/, /warā<sup>1</sup> saan<sup>5</sup>/, /ʔeekā<sup>2</sup> saan<sup>5</sup>/
- /lem<sup>3</sup>/<sub>2</sub>: /samu<sup>2</sup>t/, /nan<sup>5</sup>sy<sup>5</sup>/, /samu<sup>2</sup>t banty<sup>4</sup>k/, /nitayā<sup>2</sup> saan<sup>5</sup>/
- /phyyn<sup>5</sup>/: /syā<sup>2</sup>/, /phrom<sup>1</sup>/, /maan<sup>3</sup>/, /naŋ<sup>5</sup> sa<sup>2</sup>t/, /phaa<sup>3</sup>/ (and its compound nouns, e.g. /phaa<sup>3</sup> ʔoo<sup>3</sup>m/, /phaa<sup>3</sup> tche<sup>4</sup>t tuā<sup>1</sup>/ etc.)
- /phæŋ<sup>2</sup>/<sub>2</sub>: /kra-daan<sup>1</sup>/, /saŋ-ka-si<sup>5</sup>l/, /kra-byaŋ<sup>3</sup>/, /kha-nom<sup>5</sup> paŋ<sup>1</sup>/, /jaaŋ<sup>1</sup>/, /naŋ<sup>5</sup>/, /kra-daat<sup>2</sup>/, /ma-mua<sup>3</sup>ŋ kuan<sup>1</sup>/
- /ba<sup>1</sup>j/<sub>1</sub>: /ruu<sup>3</sup>/, /thon-nā<sup>4</sup> ba<sup>2</sup>t/, /pha<sup>3</sup>j/, /ba<sup>2</sup>t/, /tuā<sup>5</sup>/

The features [+quadrilateral], [+cubical], [+bound], [+pliant] need some qualification here. [+quadrilateral] only implies that the object has breadth and length while [+cubical] does not mean that it is actually cubical, only that it has thickness added to its length and breadth. [+bound] here implies that the object has leaves which are supposed to be (bound) together; hence /nan-sy<sup>5</sup> phi<sup>1</sup>m/ 'newspaper' is [+bound] although it is actually not bound. [+pliant] means 'can be folded' as well as 'supposed to be folded' hence /ruu<sup>3</sup>/ 'a photograph' is [-pliant] because it is not supposed to be folded.

Since the diagrams used in this paper are self-explanatory, there will be notes only for items which require explanation or modification.

As /thæŋ<sup>3</sup>/<sub>1</sub> and /kooŋ<sup>3</sup>/ are distinctive by reason of the feature [smooth], one may argue that this is not the case with /nam-khæŋ<sup>5</sup>/ 'ice' and /sa-bu<sup>2</sup>/ 'soap' which can be smooth but have the feature [-smooth]. But /naam<sup>4</sup> khæŋ<sup>5</sup>/ is quite uneven when compared with the nouns which use /thæŋ<sup>3</sup>/ like /jaaŋ<sup>1</sup> loŋ<sup>4</sup>/ 'rubber (eraser)', and /thoŋ<sup>1</sup>/ 'ingot'. /sa-bu<sup>2</sup>/ is a loanword and it would be quite legitimate to leave it out in the analysis. However, it is probably better to treat them as exceptions.

/kra-tco<sup>2</sup>ŋaaw<sup>1</sup>/ 'mirror' also has the feature [+penetrable]. This is because I have taken the feature [+penetrable] to mean 'any kind of penetration beyond the partition' which, in the case of a mirror, one achieves through vision.

It should be noted here that /le<sup>3</sup>m/<sub>2</sub> is a different classifier from the homophonous linear classifier /le<sup>3</sup>m/<sub>1</sub> in the previous section.

/kra-daa<sup>2</sup>/ 'paper' is another exception as it has the feature [+brittle]. This is a loanword from Portuguese. A colleague of mine has given an interesting explanation that it is quite possible that the early form of paper could be thick and crispy and this would have fitted the feature. However, this is merely conjecture. /jaa<sup>1</sup>h/ 'rubber mat', /na<sup>5</sup>h/ 'dried skin' and /ma-mua<sup>3</sup>h kua<sup>1</sup>h/ 'dried mango paste' are no exception if one extends the feature [+pliant] to include objects which are not supposed to be folded.

### 2.3. Round Classifiers



Nouns used for each classifier:

/lɔɔt <sup>2</sup> /:	/lɔɔt <sup>2</sup> faj <sup>1</sup> /, /lɔɔt <sup>2</sup> kəəw <sup>3</sup> /
/luuk <sup>3</sup> /, /phon <sup>5</sup> /:	/luuk <sup>3</sup> bɔɔn <sup>1</sup> /, /luuk <sup>3</sup> then-ni <sup>4</sup> /, /luuk <sup>3</sup> piŋ-pɔɔn <sup>1</sup> /, /luuk <sup>3</sup> kɔɔf <sup>4</sup> /
/huā <sup>5</sup> / <sub>1</sub> :	/huā <sup>5</sup> hɔɔm <sup>5</sup> /, /kra-thiam <sup>1</sup> /, /huā <sup>5</sup> tchaj <sup>1</sup> thaaw <sup>4</sup> /, /phak <sup>2</sup> kaat <sup>2</sup> /, /khæə-rɔɔt <sup>2</sup> /
/huā <sup>5</sup> / <sub>2</sub> :	/kra-lam <sup>2</sup> plii <sup>1</sup> /, /phak <sup>2</sup> kaat <sup>2</sup> hɔɔm <sup>5</sup> /, /huā <sup>5</sup> plii <sup>1</sup> /
/phon <sup>5</sup> / or /baj <sup>1</sup> / <sub>2</sub> :	/tæəŋ <sup>1</sup> moɔ <sup>1</sup> /, /som <sup>3</sup> /, /ma-naaw <sup>1</sup> /, /som <sup>3</sup> ʔoɔ <sup>1</sup> /, /ma-muan <sup>3</sup> /, /malakɔɔ <sup>1</sup> /, /klua <sup>3</sup> /, /tæəŋ <sup>1</sup> kwaā <sup>1</sup> /, /fak <sup>4</sup> /
/paak <sup>2</sup> /:	/hæə <sup>5</sup> /, /ʔua <sup>1</sup> /
/woŋ <sup>1</sup> /:	/woŋ <sup>1</sup> klom <sup>1</sup> /, /wæəŋ <sup>5</sup> /, /kam-laj <sup>1</sup> /, /pha <sup>3</sup> /, /maak <sup>2</sup> ruk <sup>4</sup> /
/phuā <sup>1</sup> / <sub>1</sub> :	/maa-laj <sup>1</sup> /, /naŋ <sup>5</sup> sa-tik <sup>4</sup> /
/rian <sup>5</sup> /:	/ŋya <sup>1</sup> /, /traā <sup>1</sup> /
/duaŋ <sup>1</sup> / <sub>1</sub> :	/ta-wa <sup>1</sup> /, /dya <sup>1</sup> /, /daaw <sup>1</sup> /, /win-jaan <sup>1</sup> /, /rɔɔj <sup>1</sup> pya <sup>3</sup> /
/dɔɔk <sup>2</sup> / <sub>1</sub> :	/dɔɔk <sup>2</sup> maj <sup>4</sup> /, /thɔɔŋ <sup>1</sup> jip <sup>2</sup> /, /kun-tcæə <sup>1</sup> /

The feature [+spherical] is not to be taken literally. Here, it means any round object with a third dimension. [+opening] is used for any object with a circumferential top and another part attached to it.

/luuk<sup>3</sup>/ is used for a hollow spherical object. However, its spherical feature is very strong and its use is also extended to objects which are not hollow such as all round fruit and vegetables. Hence it is used alternatively with /baj<sup>1</sup>/. This connection with fruit and vegetables results in the use of these two classifiers for any fruit and vegetable which in fact of cylindrical shape like /fak<sup>4</sup>/ 'marrow', and /tæəŋ<sup>1</sup> kwaā<sup>1</sup>/ 'cucumber'. /baj<sup>1</sup>/<sub>2</sub> here is homophonous with /baj<sup>1</sup>/<sub>1</sub> which is used for non-brittle flat quadrilateral objects.

/huā<sup>5</sup>/<sub>1</sub> which is used for a round root is likewise used for any root which may not be spherical in shape, like a carrot or a turnip.

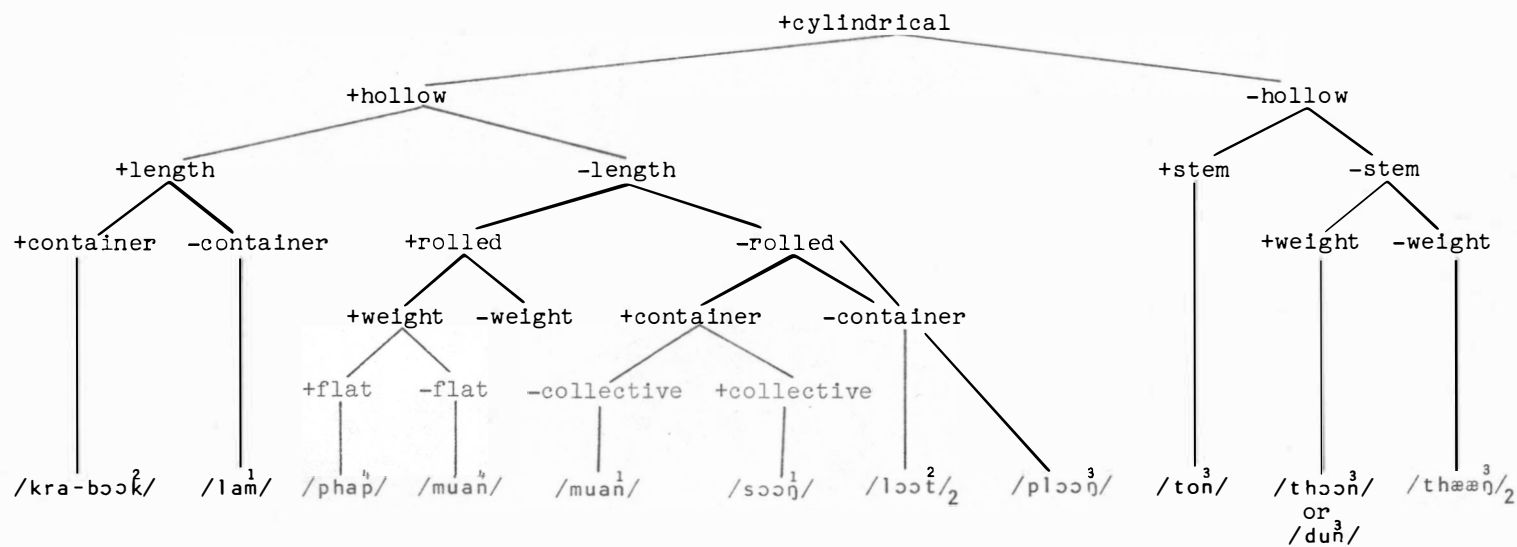
The distinctive feature for /duaŋ<sup>1</sup>/ and /dɔɔk<sup>2</sup>/ is [radiate]. It seems strange that /duaŋ<sup>1</sup>/ should be used with /win<sup>1</sup> jaan<sup>1</sup>/ 'soul' which is an abstract noun. But the Thai believe that the soul is a luminous ball which leaves a body on one's death, as often mentioned in Thai literature. Another strange noun to use with /duaŋ<sup>1</sup>/ is /rɔɔj<sup>1</sup> pya<sup>3</sup>/ 'stain'

but a stain to be used with this classifier has to be of the kind which starts with a spot and then radiates in all directions.

The use of /woŋ<sup>1</sup>/ is also extended to people sitting for a game of cards or chess; it is also used for an orchestra or a drama group, all of which are abstract. However there seems to be a concrete connection if one takes it to refer to the gathering itself.

#### 2.4. Cylindrical Classifiers





Nouns used with each classifier:

- /kra-bɔɔk<sup>2</sup>/: /pyɲ<sup>1</sup>/, /fa<sup>1</sup> tɕaa<sup>5</sup>j/, /khaa<sup>3</sup> laa<sup>5</sup>m/  
 /la<sup>1</sup>m/: /ma<sup>4</sup>j pha<sup>2</sup>j/  
 /pha<sup>4</sup>p/: /pha<sup>3</sup>a/  
 /mua<sup>4</sup>n/: /kra-da<sup>2</sup>at/, /sa<sup>5</sup>m li<sup>1</sup>i/, /kra-da<sup>1</sup>at tɕha<sup>4</sup>m-ra<sup>4</sup>/  
 /mua<sup>1</sup>n/: /bu-ri<sup>2</sup>i/, /thɔɔ<sup>1</sup>ŋ mua<sup>4</sup>n/  
 /lɔɔt<sup>2</sup>/<sub>2</sub>: /lɔɔt<sup>2</sup> kaa-fæ<sup>1</sup>æ/, /lɔɔt<sup>3</sup> ja<sup>1</sup>a/, /lɔɔt<sup>2</sup> (phom<sup>5</sup>)/, /lɔɔt<sup>2</sup> nii-ʔɔɔn<sup>1</sup>/,  
 /ro<sup>1</sup>ɔ mua<sup>4</sup>n phom<sup>5</sup>/  
 /plɔɔ<sup>3</sup>ŋ/: /ja<sup>3</sup>a/  
 /to<sup>3</sup>n/: /su<sup>1</sup>ŋ/, /sa<sup>5</sup>w/, /to<sup>3</sup>n maa<sup>4</sup>j/  
 /thɔɔ<sup>3</sup>n/: /maa<sup>4</sup>j/, /kra-bɔɔ<sup>1</sup>ŋ/, /le<sup>2</sup>k/, /pla<sup>1</sup>a/, /tɕhyak<sup>3</sup>/  
 /du<sup>3</sup>n/: /fyɲ<sup>1</sup>/, /le<sup>2</sup>k/, /maa<sup>4</sup>j/  
 /sɔɔ<sup>1</sup>ŋ/: /bu-ri<sup>2</sup>i/, /sɔɔ<sup>1</sup>ŋ tɕot<sup>2</sup> maa<sup>5</sup>j/  
 /thæ<sup>3</sup>ŋ/<sub>2</sub>: /tɕhɔɔ<sup>4</sup>k/, /din-sɔ<sup>5</sup>/, /paak-ka<sup>1</sup>a/, /phu<sup>1</sup>u-ka<sup>1</sup>n/, /ja<sup>1</sup>aŋ lo<sup>4</sup>p/,  
 /ʔaj-tim<sup>1</sup>/

/kra-bɔɔk<sup>2</sup>/ is used for /pyɲ<sup>1</sup>/ 'gun' and /fa<sup>1</sup> tɕaa<sup>5</sup>j/ 'torch' as the emphasis is on the tubular nature of the two objects. One may raise the question that /bu-ri<sup>2</sup>i/ 'cigarette' is not hollow but it is so featured because of the need to differentiate nouns used with /mua<sup>4</sup>n/ - that is rolled heavy objects - from the light ones. These two classifiers should be treated together as they derive from the verb /mua<sup>4</sup>n/ 'to roll'. It can also be argued that a cigarette is never tightly packed. /sɔɔ<sup>1</sup>ŋ tɕot<sup>2</sup> maa<sup>5</sup>j/ 'envelope' has the feature [+hollow] as it is hollow when opened.

/thɔɔ<sup>3</sup>n/ and /du<sup>3</sup>n/ are used alternatively but in applying one of these classifiers to the same object, one is more conscious of the weight when /du<sup>3</sup>n/ is used.

In this paper, I have used 28 features to differentiate 43 classifiers which are used with 173 nouns. The number of the nouns is in fact greater as some of them are generic (like 'flowers', 'trees', 'fruit'). A great number of compounds can also be formed with many nouns cited.

It may not seem economical but one should take into account that when other classifiers are analysed, the same features can be applied again. Nonetheless, it should be seen clearly from this illustration that there is a definite semantic connection between the noun and the classifier. In completing this paper, I have accidentally come across the third scale of correspondence which is the connection between the vowel of the classifiers and their shapes for all round classifiers except three (/pha<sup>4</sup>p/, /paak<sup>2</sup>/, /thæ<sup>3</sup>ŋ/) have rounded vowels while the other two groups have unrounded vowels. It may not be far from the truth to say that although language is arbitrary, there is evidence of its imitative nature which is found more in some languages than in others.

#### N O T E S

1. In this paper, J.G. Harris' device for marking tones is used, i.e. the superscript numbers; <sup>1</sup> for the mid tone, <sup>2</sup> for the low tone, <sup>3</sup> for the falling tone, <sup>4</sup> for the high tone and <sup>5</sup> for the rising tone.

/y/ will also be used for [w] and /tc/, /tch/ for /tɕ/, /tɕh/ respectively.

2. My apologies for not giving an English translation for the non-Thai reader as it tends to distract one's attention in reading. A list of English equivalents is given, in the same order as the words occur in the article, in the appendix.

## APPENDIX

### 2.1. Nouns used with linear classifiers

- /lem<sup>1</sup>/₁: *knife, hoe, spade, needle, scissors, candle, oar, axe, prongs, spatula.*
- /daam<sup>3</sup>/: *fan, pen, a kind of Chinese weapon.*
- /khañ<sup>1</sup>/: *bow, bow, umbrella, fishing rod, spoon, fork, ladle, banjo.*
- /saa<sup>5</sup>/: *road, way, river, stream, canal.*
- /señ<sup>3</sup>/: *fan belt, egg noodle, transparent noodle, silk thread, cotton thread, belt, shoe laces, hair, wire, thread, string, steel wire, telephone wire, electrical wire.*

### 2.2. Nouns used with quadrilateral classifiers

- /thæñ<sup>3</sup>/₁: *iron, ingot, rubber eraser.*
- /kooñ<sup>3</sup>/: *stone, clay, shit, cloud, a loaf of bread, brick, ice cube, soap, sugar cube, plasticine.*
- /baañ<sup>1</sup>/: *door, window, mirror.*
- /daañ<sup>3</sup>/: *wall, fence, room wall, side (e.g. of a triangle).*
- /tcha-bap<sup>2</sup>/: *letter, newspaper, journal, document.*
- /lem<sup>3</sup>/₂: *exercise book, book, diary, magazine.*
- /phyñ<sup>5</sup>/: *mat, carpet, curtain, hide, cloth (and its compound nouns, e.g. nappy, towel, etc.)*
- /phañ<sup>2</sup>/₂: *plank, corrugated iron, tile, a slice of bread, rubber mat (in the car), dried skin, paper, dried mango paste.*
- /ba<sup>1</sup>/₁: *photograph, banknote, card (for playing), (invitation) card, ticket.*

## 2.3. Nouns used with round classifiers

/lɔɔt <sup>2</sup> /:	<i>electric bulb, glass bulb (used for experiment).</i>
/luuk <sup>3</sup> /, /phon <sup>5</sup> /:	<i>ball, tennis ball, pingpong ball, golf ball.</i>
/hua <sup>5</sup> / <sub>1</sub> :	<i>onion, garlic, turnip, long lettuce, carrot.</i>
/hua <sup>5</sup> / <sub>2</sub> :	<i>cabbage, lettuce, banana flower.</i>
/phon <sup>5</sup> /, /baj <sup>1</sup> / <sub>2</sub> :	<i>melon, orange, lemon, pomelo, mango, papaya, banana, cucumber, marrow.</i>
/paak <sup>2</sup> /:	<i>fishing net, fishing net.</i>
/won <sup>1</sup> /:	<i>circle, ring, bracelet, card game, chess game.</i>
/phuah <sup>1</sup> /:	<i>garland, a bunch of rubber bands.</i>
/rian <sup>5</sup> /:	<i>coin, medal.</i>
/dua <sup>1</sup> /:	<i>sun, moon, star, soul, stain.</i>
/dɔɔk <sup>2</sup> /:	<i>flowers, a kind of Thai sweet, key.</i>

## 2.4. Nouns used with cylindrical classifiers

/kra-bɔɔk <sup>2</sup> /:	<i>pistol, torch, Thai dessert contained in bamboo tube.</i>
/lam <sup>1</sup> /:	<i>bamboo stem.</i>
/pha <sup>4</sup> p/:	<i>cloth (rolled).</i>
/muan <sup>4</sup> /:	<i>(rolled) paper, a roll of cotton wool, toilet roll.</i>
/muan <sup>1</sup> /:	<i>cigarette, rolled pancake.</i>
/lɔɔt <sup>2</sup> / <sub>2</sub> :	<i>drinking straw, syringe, a roll of hair, fluorescent tube, hair roll.</i>
/plɔɔŋ <sup>3</sup> /:	<i>reed.</i>
/ton <sup>3</sup> /:	<i>log, pole, trees.</i>
/thɔɔn <sup>3</sup> /:	<i>wood, a weapon made of wood, iron bar, a big piece of fish, rope.</i>
/du <sup>3</sup> /:	<i>log (for fire), iron bar, wood.</i>
/sɔɔŋ <sup>1</sup> /:	<i>a pack of cigarettes, envelope.</i>
/thæŋ <sup>3</sup> / <sub>2</sub> :	<i>chalk, pencil, pen, painting brush, rubber eraser, ice lolly.</i>

BIBLIOGRAPHY

NOSS, Richard B.

1964 *Thai Reference Grammar*. Washington, D.C.: Foreign Service  
Institute.

WAROTAMASIKKHADIT, Udom

1963 *Thai Syntax: An Outline*. Doctoral dissertation, University  
of Texas, Austin.

## HOW MUCH IS ENGLISH INFLUENCING THE LANGUAGE OF THE EDUCATED BANGKOK THAIS?

WILAIWAN KANITTANAN

The Bangkok way of life has been to a certain degree Westernised or to be more particular it has been Americanised. Hot dogs, hamburgers, ice-cream cones, blue jeans, American movies, the latest popular songs from Europe and America are prevalent in Bangkok. American television programs like 'Truth or Consequences' or 'The Newly Wed Game' are imitated on Thai television. Many teenagers and educated people are living an American way of life in Bangkok.

Not only have they been living an American or Western-style life, they have brought many English features into their language. Many words, phrases, and sentences used by educated people, especially people educated overseas are often esoteric. One often hears remarks such as, "You almost have to be able to speak English to understand this Thai article." This sounds peculiar but this is what is happening in the daily usage of the Thai language in Bangkok by educated Thais.

Lots of English words are imported into educated Bangkok Thai. They are used as part of the daily vocabulary. Not that Thai lacks the words the educated Thais want to use. Native words are thought of as not being as expressive as the English ones. Words that are used quite frequently include tax, class, test, progress, service charge, lecture, nominate, private, semester, part time, advisor, and shake hands.

People educated overseas sometimes have difficulty making themselves understood because they have a different accent. They sometimes have to repeat what they have just said twice or three times before they can get a soft drink they order in an ordinary coffee shop. The names of the soft drinks that they have most difficulty with are 'Sprite' and 'Green Spot'. The normal local people pronounce them [sa-páy] and [kiyn-sa-pòt] thus they become baffled when they hear the original English pronunciation.

Not only do the people educated overseas have an incomprehensible pronunciation, they also used Thai words which an average person cannot understand. Words such as

- [chán] 'class'  
 [khâa-ní-yom] 'value'  
 [chôn-wâaŋ] 'gap' as in 'generation gap'  
 [koo-ra-nii-sák-săa] 'case study'  
 [nûay-kít] 'credit' as used in universities

are all translated from English into Thai and these words are used only among educated people.

Very often, the people educated overseas would use phrases or sentences that a native Thai would not use, for example:

- (1) khòp-khun sām-ráp ʔaa-haǎn-yen  
 thank you for dinner  
 'Thank you for dinner.'
- (2) nân pen kham-thăam thîi dii  
 that is question which good  
 'That is a good question.'
- (3) phǒm dii-cay thîi khun thăam  
 I glad that you ask  
 'I am glad that you asked.'
- (4) nân mây leew læy  
 that not bad at all  
 'That is not bad at all.'

An average Thai would say only [khòp-khun] 'thank you' or [khòp-khun mâak khráp] 'thank you very much' with a polite particle for male speaker. And he would say nothing in the second and the third cases. He would go ahead and answer the question he is being asked without making any comment on the question and the questioner. In the fourth case, he would make a more particular comment which could be translated as 'it's beautiful', 'it's delicious', etc. The expressions in the examples above are definitely translated from English.

Many new sentence structures have been brought into use in educated Bangkok Thai from English. The two most popular ones are as follows:

#### The Passive Voice Sentence Type

phôn-naan khòŋ-khăw thûuk khát-lâak pay sa-dxxŋ thîi naan  
 work his passive choose go show at fair  
 marker?

'His work was chosen to be displayed at the fair.'



khaw thùuk lîak pen hũa-nâa khon-naan  
 he passive choose be leader worker  
 marker?

*'He was chosen to be the leader of the workers.'*

khaw thùuk chom tồ-nâa  
 he passive praise in front  
 marker?

*'He was praised to the face.'*

The word [thùuk] is taken to be a passive marker by most Thai students and grammarians. That is why the people educated overseas use it whenever they want to make a passive voice sentence which sounds awkward to any native Thai. But the Thai language does not have passive voice sentence type in the same sense as the English language does. That is in Thai not all active voice sentences with transitive verbs can be changed into passive voice sentences. For example, in Thai one can say,

khăw hễn mxxw  
 he see cat

*'He saw the cat.'*

and

khăw kin ma-mûaŋ  
 he eat mango

*'He ate the mango.'*

but not

\*mxxw thùuk khăw hễn  
 cat passive he see  
 marker?

*'The cat was seen by him.'*

\*ma-mûaŋ thùuk khăw kin  
 mango passive he eat  
 marker

*'The mango was eaten by him.'*

As Professor Pongsri Lekawatana pointed out in her article "The so-called Passive in Thai", that [thùuk] and [dooy] are verbs not passive voice markers. As verbs they mean to suffer or to experience something unpleasant. Thus they can be used only with a subject that suffers or experiences something unpleasant. They cannot be used with animate subject, e.g. rulers, stones. When the educated people use the verbs [thùuk] and [dooy], they violate all these restrictions. They use them in sentences in which [thùuk] and [dooy] do not convey unpleasant experiences such as:

khaw thùuk chœn  
 he passive invite  
 marker

*'He was invited.'*

nǎŋ-sǎi khaw thòuk klàaw thǎŋ yàaŋ-mâak  
*book his passive talk about a lot*  
 marker

*'His book is talked about a lot.'*

Though this type of sentence is understandable, it is peculiar to an average Thai who does not speak or know English.

#### The 'It is ...' Type

This type of sentence is so widely used among the educated and students now that it almost sounds like a native sentence structure. But it is this type of sentence that those who do not speak English get confused with most. For example:

- (1) man pen khwaam-khít khǎŋ-khǎw thǐi ca pay  
*it be idea his which will go*

*'It is his idea to go.'*

- (2) man pen rǎaŋ pra-laàt thǐi khǎw ca maa  
*it be matter strange which he will come*

*'It is strange that he will come.'*

- (3) man pen khwaam-rák khǎŋ-rǎxx khǎŋ-khǎw  
*it be love first his*

*'It is his first love.'*

Those who do not know English would use the following sentences instead:

- (1) khǎw khít ca pay  
*he think will go*  
*'He thinks of going.'* (?)

- (2) pra-laàt khǎw ca maa  
*strange he will come*

- (3) khwaam-rák khǎŋ-rǎxx khǎŋ-khǎw  
*love first his*

For them, [man] 'it' is normally used as a third person singular pronoun. When it is not used as a person pronoun, it does not occur with the verb 'to be' [pen], for example [man cǎp] 'it hurts'.

I once read two passages written by a teacher and a student educated overseas to my relatives who live in a province next to Bangkok in order to find out how much they could understand. The passages were about the life of university students and the life of Bangkok people. The passages contained translated words and translated sentence types as described above. The relatives were common rice farmers who could read and write. They were able to understand about half of what the passages were describing and that could be partly guess work too.

This could be a universal trend in developing countries where people go to English-speaking countries to get educated. When they come back they bring many things with them, among them a different dialect of their own language. Since they are those who do more writing work than average people they impose changes on their language in the long run.



# FORMS AND MEANINGS OF THE THAI PARTICLE SI

JOSEPH R. COOKE

## 0. SUMMARY OF ARTICLE

The Thai form *si* is a discourse particle having various pronunciations and used in a wide variety of ways. The variant pronunciations include /si/, /sii/, /sɨ/, /sʰi/, /sʰi/, and (for some speakers) /sʰi/; and the varying usages include action-inducement utterances (commands, suggestions, invitations, requests), responses to questions and to question-raising statements, inferential comments, and statements noting new information. All these forms and usages have one meaning in common - that of signalling a logical, necessary, or expectable response. And then the variations in form signal further distinctions as follows: /si/ or /sii/ for non-involvement, /sɨ/ for definiteness, /sʰi/ for persuasion, /sʰi/ for personal need or wish, and /sʰi/ for personal wish plus persuasion. Under certain circumstances these variants may be neutralised to /si/; and the forms /sɨ/ and /sʰi/ may be raised to signal intensification of meaning. The above phenomena are exemplified in this paper through the presentation of a wide range of data; and the data are then accounted for by means of relevant explanations and generalisations.

## 1. INTRODUCTION

### 1.1. SI AND THE CLASS OF DISCOURSE PARTICLES

The form *si*, with its various pronunciations and meanings, comprises one of a class of forms in Thai sometimes designated as sentence-final particles but perhaps more appropriately identified as discourse particles. These particles usually but not always occur at the ends of sentences, and they generally signal various types of commands, questions, responses, statements, etc. They also constitute links of

various kinds with the linguistic and non-linguistic context of the discourse or linguistic interchange within which they occur.

## 1.2. SOURCES OF INFORMATION

Some of these particles prove extraordinarily resistant to definition, analysis, or explanation. For one thing, some occur with a variety of pronunciations the precise significance of which is extremely difficult to determine. And certain particles are used in such a variety of ways that one is hard put to it to discover what function they could possibly have. Then again the line between what is acceptable or grammatical and what is not sometimes seems so tortuous and arbitrary that one wonders how a native speaker ever learns to use the forms correctly or to understand the usage of other speakers.

This paper constitutes an attempt to make sense out of the bewildering ins and outs of the forms and meanings of just one of these particles - the form *si*. I have selected this particular form for consideration because it has been, for me, the most bewildering and complex of the lot. Also, I am hopeful that light shed in this area may lead to a more insightful exploration of a wider range of phenomena relating to the whole class of discourse particles.

In preparing this paper, I have, of course, had access to a body of published material (see bibliography); but most of the information contained herein has been obtained from several years of intermittent bedeviling of a number of very patient and helpful native speakers. The following have all assisted me by spending considerable amounts of time sharing their knowledge and understanding with me: Dr Prapin Manomaivibool, Ms Nisa Udomphol (now Ms Sakdechayont), Ms Peansiri Ekniyom, Mr Chare Vathanaprida, Ms Subhaphorn Vathanaprida, Ms Pimpun Suwanamalik (now Ms Fitzpatrick), Ms Niphapharn Chutrakul, Dr Navavan Bandhumedha, and Ms Arada Kiranand. I have also leaned very heavily upon an unpublished paper prepared for me by Ms Udomphol, entitled 'Semantic Functions of the Thai Particle /Si/'. A number of examples cited below have been taken from her work.

In general, the data and explanations which follow are presented in terms of the usage of my most recent informant, Ms Kiranand. Other speakers will certainly differ from Ms Kiranand in their use of *si*, and some of these differences have been recorded in my notes; but many other differences assuredly have not, for some of my data were gathered at a time when my perceptions and understanding were more limited than now. Also, unfortunately, I no longer have access to my original sources of information, so I cannot check my data in the light of more recent

insights. In any case, my presentation is structured around Ms Kiranand's speech patterns. Significant variations from those patterns will be pointed out where relevant.

### 1.3. FORMS AND MEANINGS OF SI AND THE TREATMENT OF THEM HERE PROPOSED

Now to an examination of the particle si. And in order to lay a foundation for our discussion, I must explain that si occurs with the following forms: /sɪ/, sɪ̃/, sɪ̃/, sɪi/, sɪ̃i/, and for some speakers, /sɪ̃i/.<sup>1</sup> All of these forms seem to possess some element of meaning which they hold in common; yet each can, for the most part, be differentiated from the others by some distinct and consistent meaning that it possesses. I shall attempt in this paper to identify the basic meaning common to all forms, and to isolate the meanings that distinguish each variant from the others. As I do this, it will soon become evident that the bulk of the paper is concerned with semantic problems; and my approach in dealing with these is first to present data, then to formulate hypotheses, and then in certain cases to show how these hypotheses apply. The paper then concludes with a summary of my conclusions and a couple of suggestions concerning possible future research.

## 2. THE BASIC MEANING OF SI AS EXEMPLIFIED BY THE VARIANT /sɪ̃/

Let us first take up the matter of the basic meaning that is characteristic (as I suggest) of si in all its varied forms and occurrences. And, in order to bring the wealth of data down to manageable proportions, let me present a number of examples of just one of the variant forms, namely /sɪ̃/. I choose this particular form because it happens to occur in a rather wide range of situations; and, once such occurrences are explained, we will find we have a convenient basis for going on to account for the other variants.

### 2.1. EXAMPLES OF TYPES OF OCCURRENCE OF /sɪ̃/

The following examples are arranged according to varied categories of occurrence or usage: commands, suggestions, invitations, requests, responses to questions and to question-raising statements, inferential statements, and utterances noting new information. These categories should not, however, be taken too seriously, for they merely provide a convenient means for setting forth the data. When a given example fits into one category or another is not a matter of crucial importance. The point is that /sɪ̃/ occurs in each of the varied contexts, and we must find some account of its meaning that is consonant with this wide variety of occurrences.

Square brackets, below, mark information as to possible situations in which the utterance in question might occur.

### 2.1.1. Action-inducement Utterances

These comprise various kinds of utterances in which the speaker is prompting the addressee to some particular action. They include commands (see examples 1 and 2 below), suggestions (3-7), invitations (8, 9), and requests (10-12).

- (1) /pəət<sub>1</sub> pratuu<sub>2</sub> sî/ 'Open<sub>1</sub> the door<sub>2</sub>.' [It's time for the store to open, and it is the addressee's responsibility to perform this duty. Or: A third party's hands are full, and he can't open the door himself, but the addressee is there handy to help him. Or: The addressee appears to be uncomfortable sitting in a stuffy, closed room.]
- (2) /yàa<sub>1</sub> khàp<sub>2</sub> rew<sub>3</sub> !sî/ 'Don't<sub>1</sub> drive<sub>2</sub> so fast<sub>3</sub>.' [The speaker thinks the addressee is driving too fast.] (The exclamation symbol here, and in example 18 below, indicates an emphatic raising of the pitch of /sî/.)
- (3) /faŋ<sub>1</sub> sî, phrəʔ dii<sub>2</sub>/ 'Listen<sub>1</sub>! (That's) beautiful<sub>2</sub>.' [The speaker hears some beautiful music and calls it to the attention of the addressee.]
- (4) /khīan<sub>1</sub> hây dii<sub>2</sub> sî, lɛɛw<sub>3</sub> ca<sub>4</sub> dâay<sub>5</sub> raanwan<sub>6</sub>/ 'Write<sub>1</sub> nicely<sub>2</sub> now, and<sub>3</sub> (you)'ll<sub>4</sub> get<sub>5</sub> a reward<sub>6</sub>.' [A mother wants her child to write to his grandfather, and she offers him a reward if he writes a nice, neat letter.]
- (5) /səu<sub>1</sub> sâa<sub>2</sub> tua nân<sub>3</sub> sî, sūay<sub>4</sub> dii<sub>5</sub>/ '(Why don't you) buy<sub>1</sub> that<sub>3</sub> shirt<sub>2</sub>? It's nice<sub>5</sub> and pretty<sub>4</sub>.'
- (6) /khăw<sub>1</sub> hây<sub>2</sub> kô<sub>3</sub> ʔaw<sub>4</sub> sî/ 'He<sub>1</sub>'s giving<sub>2</sub> (it to you), so<sub>3</sub> take<sub>4</sub> (it).' [The speaker is encouraging the addressee to accept the offer being extended to him.]
- (7) /kô<sub>1</sub> yàa<sub>2</sub> nân<sub>3</sub> sî/ 'Well then<sub>1</sub> don't<sub>2</sub> sit<sub>3</sub> (there).' [The addressee has just indicated verbally that he is reluctant to seat himself. Perhaps he is afraid the chair won't take his weight, or he has noticed something spilled on it.]
- (8) /khâw maa<sub>1</sub> sî/ '(Do) come in<sub>1</sub>.' [The speaker is welcoming someone at the door.]



- (9) /kin khâaw<sub>1</sub> kòɔn<sub>2</sub> sî, lɛ̃ɛw<sub>3</sub> khây pay<sub>4</sub>/ 'Eat<sub>1</sub> first<sub>2</sub>, then<sub>3</sub> (you can) go<sub>4</sub>.' [The addressee is about to leave; but it is almost time to eat, so the speaker urges him to stay for the meal.]
- (10) /yìp<sub>1</sub> dɪnsɔ̌ɔ<sub>2</sub> hây<sub>3</sub> nòy<sub>4</sub> sî/ 'Hand me the pencil, (would you?)' (grasp<sub>1</sub> pencil<sub>2</sub> for (me)<sub>3</sub> a little<sub>4</sub>) [The pencil is within easy reach of the addressee, and the speaker cannot conveniently reach it for himself.]
- (11) /khɔ̌ɔ<sub>1</sub> nân<sub>2</sub> dūay<sub>3</sub> khon<sub>4</sub> sî/ 'May I join you? (request<sub>1</sub> sit<sub>2</sub> with<sub>3</sub> (you) (one) person<sub>4</sub>) [Speaker asks permission to join and sit down with a group of his friends.]
- (12) /tham<sub>1</sub> hây<sub>2</sub> thi<sub>3</sub> sî/ 'Would you do it for me?' (do<sub>1</sub> for<sub>2</sub> (me) (one) time<sub>3</sub>) [The speaker asks the addressee to do some small task for him.]

### 2.1.2. Answers to Questions

These include answers to yes-or-no questions (13-15 below) and to content questions asking who? what? when? where? etc. (16, 17). In the following examples, S1 and S2 differentiate two speakers in a given utterance-and-response interchange.

- (13) S1 /khun<sub>1</sub> khít<sub>2</sub> wâa<sub>3</sub> khâw<sub>4</sub> ca<sub>5</sub> maa<sub>6</sub> máy<sub>7</sub>/ S2 /maa<sub>6</sub> sî/ 'Do you<sub>1</sub> think<sub>2</sub> that<sub>3</sub> he<sub>4</sub>'ll come<sub>6</sub>? 'Sure he will.' (The form /mây/ signals a question that calls for a yes-or-no answer) [The second speaker has perhaps just talked to the third party on the phone and so knows he is coming. Or: the third party had promised to come, and the second speaker knows his promises are reliable.]
- (14) S1 /khun<sub>1</sub> khít<sub>2</sub> wâa<sub>3</sub> fôn<sub>4</sub> ca<sub>5</sub> mây<sub>6</sub> tòk<sub>7</sub> lǎə<sub>8</sub>/ S2 /kô mây<sub>6</sub> tòk<sub>7</sub> na sî/<sup>2</sup> 'You think it won't rain?' 'Of course it won't' (you<sub>1</sub> think<sub>2</sub> that<sub>3</sub> rain<sub>4</sub> will<sub>5</sub> not<sub>6</sub> fall<sub>7</sub>?<sub>8</sub>) (The form /lǎə/ signals a yes-or-no question where the speaker has received some clue as to the addressee's expected response; the sequence /kô ... na/, here and below, conveys the idea that the speaker is stating something that he feels should be obvious to the addressee.)
- (15) S1 /khun<sub>1</sub> ca<sub>2</sub> mây<sub>3</sub> klàp bân<sub>4</sub> lǎə<sub>5</sub>/ S2 /klàp<sub>4</sub> sî. thammay<sub>6</sub> ca<sub>2</sub> mây<sub>3</sub> klàp<sub>4</sub>/ 'Aren't<sub>3</sub> you<sub>1</sub> going to<sub>2</sub> go home<sub>4</sub>? 'Sure I am. Why<sub>6</sub> would<sub>2</sub>n't<sub>3</sub> (I)?'.
- (16) S1 /thæ<sub>1</sub> ca<sub>2</sub> pay<sub>3</sub> hǎa<sub>4</sub> khray<sub>5</sub>/ S2 /kô 'aacaan<sub>6</sub> na sî/ 'Who<sub>5</sub> are you<sub>1</sub> going to<sub>2</sub> go<sub>3</sub> see<sub>4</sub>? 'The teacher<sub>6</sub>, of course.' [The two speakers have been puzzling over an assignment, and the second

speaker has just previously indicated that he is going to inquire about it.]

- (17) S1 /ʔaw<sub>1</sub> kii<sub>2</sub> bàat<sub>3</sub>/ S2 /kâ sîp<sub>4</sub> bàat<sub>3</sub> na sî/ 'How much do you want? (want<sub>1</sub> how many<sub>2</sub> baht<sub>3</sub>) 'Well ten<sub>4</sub> baht<sub>3</sub>.' [The second speaker feels the answer is obvious. Perhaps the price is marked in plain view, or perhaps he has just quoted the price shortly before.]

### 2.1.3. Responses to Question-raising Statements

These are responses somewhat like answers to questions, but here the addressee has not actually asked a question. Rather, he has made a statement that raises or calls to mind a question of fact or understanding that the speaker feels requires comment. So he responds with some appropriate confirmation, correction, or explanation. Note that in certain types of such responses the particle /sî/ does not occur at the end of the sentence but after a noun phrase or subordinate clause which functions as the focus of the predication (see 20-22).

- (18) S1 /chán<sub>1</sub> waa<sub>2</sub> wanníi<sub>3</sub> fôn<sub>4</sub> thaa ca<sub>5</sub> mây<sub>6</sub> tòk<sub>7</sub>/ S2 /tòk<sub>7</sub> !sî/ 'I don't think it'll rain today.' (I<sub>1</sub> think<sub>2</sub> today<sub>3</sub> rain<sub>4</sub> apparently will<sub>5</sub> not<sub>6</sub> fall<sub>7</sub>) 'Sure it will.'
- (19) S1 /duu<sub>1</sub> sî, khaw<sub>2</sub> kamlan<sub>3</sub> láan<sub>4</sub> chaam<sub>5</sub>/ S2 /tèŋŋaan kan<sub>6</sub> léew<sub>7</sub> kô<sub>8</sub> tôn<sub>9</sub> chûay<sub>10</sub> kan<sub>11</sub> tham naan<sub>12</sub> sî/ 'Look<sub>1</sub>, he<sub>2</sub>'s<sub>3</sub> washing<sub>4</sub> the dishes<sub>5</sub>.' 'Well, they're married<sub>6</sub>, and<sub>7</sub> so<sub>8</sub> they've got to<sub>9</sub> help<sub>10</sub> each other<sub>11</sub> do the work<sub>12</sub>.'
- (20) /khon<sub>1</sub> nán<sub>2</sub> sî sūay<sub>3</sub>/ 'That's the one that's pretty.' (person<sub>1</sub> that<sub>2</sub> is pretty<sub>3</sub>) [The addressee has just expressed his opinion that some other person than the one here referred to is pretty, so the speaker here refocuses the addressee's attention on the one to whom he feels the description more fully applies.]
- (21) /fôn<sub>1</sub> yàan<sub>2</sub> nîi<sub>3</sub> sî tham hây<sub>4</sub> nám thûam<sub>5</sub> dâay<sub>6</sub> nâyŋây<sub>7</sub>/ 'This is the kind of rain that can easily cause floods.' (rain<sub>1</sub> kind<sub>2</sub> this<sub>3</sub> cause<sub>4</sub> water to overflow<sub>5</sub> can<sub>6</sub> easily<sub>7</sub>) [The addressee has just remarked about how serious the rain storm is. Or: he has just made light of the storm's importance.]
- (22) /pen<sub>1</sub> dèk<sub>2</sub> sî dii<sub>3</sub>/ 'Being<sub>1</sub> a child<sub>2</sub> is wonderful.' 'It's when you're a child that you're really well off.' [The addressee has just indicated what a wretched life children lead. Or: he has just been talking about what fun he had as a child.]

## 2.1.4. Inferential Statements

These are utterances in which the speaker draws some kind of inference from something he hears or observes.

- (23) /khun<sub>1</sub> khít waa<sub>2</sub> chán<sub>3</sub> tòk lon<sub>4</sub> sî/ 'You<sub>1</sub> must think<sub>2</sub> I<sub>3</sub> agree<sub>4</sub>.'  
[The speaker has just learned from the addressee that the latter, without consultation, has gone ahead with plans for a party to be held at the speaker's house.]
- (24) /khânñòk<sub>1</sub> thanñ<sub>2</sub> tòn nî<sub>3</sub> rôt<sub>4</sub> tít<sub>5</sub> 'iik<sub>6</sub> sî/ 'I gather there's another traffic jam outside now.' (outside<sub>1</sub> streets<sub>2</sub> now<sub>3</sub> cars<sub>4</sub> jammed<sub>5</sub> again<sub>6</sub>) [The addressee has just come in at 5:00 p.m., complaining about the difficult time he has had getting through town in his car.]
- (25) /fñn<sub>1</sub> tòk<sub>2</sub> léew<sub>3</sub> sî/ 'So it's raining now.' (rain<sub>1</sub> falls<sub>2</sub> now<sub>3</sub>)  
[The speaker doesn't have first-hand knowledge of the weather outside, but he sees the addressee come in, soaked from head to foot.]

## 2.1.5. Statements Noting New Information

These are statements made in response to some new development or fact that has just come to the attention of the speaker, or that is being brought to the attention of the addressee for the first time. Statements of this kind always have the form /léew/ 'now' as a part of the predication.

- (26) /fñn tòk léew sî/ 'Hey, it's raining!' Or: 'Aw shucks, it's raining!' [Note that this utterance, as it stands, is ambiguous. The presence or absence of disappointment will have to be determined by the speaker's tone of voice, or by his general deportment. Furthermore, only a knowledge of the situation will make clear whether the statement is a response to a new development, as here, or whether the speaker is making an inference, as in 25, above.]
- (27) /phleen<sub>1</sub> ràem<sub>2</sub> léew<sub>3</sub> sî/ 'There, the music is getting started.' (song<sub>1</sub> begins<sub>2</sub> now<sub>3</sub>) [The speaker and the addressee are at a concert and have been waiting for the music to begin.]
- (28) /khun<sub>1</sub> bunmii<sub>2</sub> ca<sub>3</sub> tènñaan<sub>4</sub> léew<sub>5</sub> sî/ 'Well, I see Mr Boonmii is getting married.' (Mr<sub>1</sub> Boonmii<sub>2</sub> will<sub>3</sub> marry<sub>4</sub> now<sub>5</sub>) [The speaker sees an article in the paper and discovers for the first time that Boonmii is getting married.]

- (29) /khun bunmii ca tènnaan léew nîi sî/ 'But Mr Boonmii is getting married now.' (The form nîi signals new information contrary to the expectations or understanding of the addressee.) [The addressee has just proposed the name of Boonmii as an officer in a singles group.]

## 2.2. ACCOUNTING FOR THE GENERAL MEANING OF /sî/ (AND SI)

The above examples should suffice to reveal something of the range of occurrences of the form /sî/. But what does the particle mean? Can there in fact be a single meaning that covers such a diversity of uses? This is the problem to which I shall now address myself.

### 2.2.1. The Meaning of /sî/ (or si) as Presented in the Literature

The most commonly proposed explanations for the meaning of si involve the idea of emphasis. So McFarland 1954:863; Thai-Thai Dictionary 1976: 911; Manitcharoen 1964:1356; Haas 1964:539; Bhamoraput 1972:24. Two of these sources, McFarland and the Thai-Thai Dictionary, indicate further that /sî/ is used to express the imperative; and Bhamoraput, in a similar vein, states that si indicates an exhortation. Then Brown (1969: 3.35) provides the more explicit information that /sî/ is used when "speaker urges hearer to do something that should obviously be done." So here again something of the idea of the imperative is conveyed.

Noss (1964:210), for his part, gives us a definition that includes both the idea (approximately) of the imperative and of emphasis. He suggests that si conveys the idea that 'this is the correct behaviour or belief (change yourself if necessary)'. Then he goes on to explain that si "is used most commonly to urge action on the part of someone who is not acting, or to change the course of action of someone who is .... A second use ... is in emphatic statements, where it either expresses or urges agreement."

Actually, none of the above explanations quite suffices to fit all the contexts in which si occurs.<sup>3</sup> The idea of emphasis seems plausible enough in some contexts, but it doesn't seem to be too relevant in the case of invitations or requests (examples 8-12 above), or in the case of inferential statements (23-25), or of statements noting new information (26-29). Similarly, the idea of the imperative appears relevant in some cases; for si certainly can be used with commands. In fact if we were to broaden the term 'imperative' to cover all the inducement-to-action utterances above (examples 1-12), we could considerably widen the applicability of this definition. However, there are other ways of giving commands that make no use of si (e.g. utterances occurring with

the particles /ná/ and /thèʔ/ and occasionally utterances with no particle at all); and there are any number of other things that can be said in order to get people to do things. And then, when all is said and done, we are still left with occurrences of si that convey neither the idea of the imperative nor the idea of emphasis.

Noss, I believe, comes close to the mark when he proposes the idea of 'correct behaviour or belief'. In fact, one might have difficulty demonstrating conclusively that this explanation falls short. However, I do think it is possible to improve on Noss's proposal; so, rather than argue the point, I should like to suggest a definition that I feel comes a little closer to accounting for the meaning and usage of this particle.

### 2.2.2. The Meaning of /sî/ (or si) as here Proposed

I suggest, then, that the particle si conveys the idea that something - i.e. the thing being commanded, requested, suggested, stated, affirmed, inferred, pointed out - is a response that in the speaker's opinion naturally, logically, expectably, assuredly follows from the situation in question. In other words, si signals the fact that a given response is obvious, expectable, or certain under the circumstances.

### 2.2.3. How the Notion of 'Expectability' Applies in Various Contexts

This idea of expectable response will serve, I believe, to explain the various types of usage to which si is subject. But, what constitutes an expectable response? And how does this idea of expectable response apply to the different kinds of utterances exemplified above?

In order to answer these questions, it is helpful to divide the various types of utterances where si occurs into two groups. The first group consists of those utterances which call for some appropriate or reasonable response on the part of the addressee (see the action-inducement utterances, as in examples 1-12). The second group consists of those which signal an expectable or assured response on the part of the speaker (as in examples 13-29).

#### 2.2.3.1. *'Expectability' in Action-inducement Utterances*

The first group of utterances, then, signal that something is to be done by the addressee; and the something, whatever it is, must be something that is expectable within the verbal or situational context within which the given si utterance occurs. Now this expectability will in some cases be self-evident, both to the speaker and to the addressee,

in the light of the situation as it stands. For example, it may be understood that it is time for the addressee to open the store; so then the situation naturally calls for the addressee to open it (see item 1 above). Or the speaker may be extending the addressee an invitation to come into his house (example 8), or to be seated; and the natural, expectable response is for the addressee to accept. In cases such as these, the situation plus the stimulus utterance in themselves provide all the grounds necessary for expecting the given response. The response is expectable without any further comment or explanation.

In other cases, the expectability of the action in question is not self-evident unless the speaker provides some explanation or points to some consideration that reinforces the expectability of the response in a given instance. For example, a speaker might see a shirt that he thinks the addressee should buy, but he cannot simply out of the blue urge the latter to buy it, using the word *si*. However, if the speaker explains that the shirt is pretty, then that explanation provides reason enough for the speaker to feel the addressee's response is expectable, and so he can appropriately use *si* in calling forth that response (example 5).

In short, a speaker may tack *si* on to an action-inducement utterance when there is something about the situation that in and of itself naturally calls for the action in question. But if the occasion doesn't speak for itself, the speaker will verbally supply information or suggest some consideration that explains why he feels the response in question is called for.

#### 2.2.3.1.1. 'Expectability' that is Self-explanatory

Situations that speak for themselves, or those in which the expectability of the called for response is self-explanatory, include those in which anyone might be expected to act in the way indicated. For example, an employee may be expected to fulfil responsibilities assigned to him (like opening the store door every morning, or sweeping every evening). A child may be expected to eat when food is set before him. A person may be expected to listen to the music at a concert, or to open a door for a friend whose hands are full, or to enter a house when he is welcomed at the door by his host, or to take steps to get warm when he is cold. All of these are things a person might be urged, told, asked, invited to do, using the particle *si*, with no further explanation. The explanation may be provided if the speaker wishes, but it need not be.

Some types of utterance in which the situation normally can be considered as self-explanatory are worthy of note here:

1) Corrective commands. These (as opposed to preventive commands, to be discussed below) are commands directed toward the addressee to get him to do something he is not doing but in the speaker's opinion should do, or to stop doing something that is contrary to the speaker's wishes or expectations. For example, a mother tells her child who is toying with his food to get busy and eat; or a passenger tells a driver not to drive so fast (item 2). Note that in situations of this sort, the addressee is not necessarily expected to already know without being told (though he may know) what action is called for or why. The command itself informs him that his present behaviour is undesirable, and that he should therefore either stop what he is doing, or start acting in a different way. In other words, the command itself reveals what the expectable response is; and the speaker, in using *si* is expressing his opinion of what is expectable.

2) Requests or invitations arising out of some present and immediate need or wish. For example, the speaker needs a pencil that is out of convenient reach, so he asks the addressee to pass it to him (item 10); or someone wants to join a group of friends, so he asks their permission to do so (11); or the addressee is standing at the door, and the host invites him in (8); or the host invites his guest to be seated. Here again, the expressed wish or invitation of the speaker provides all the information necessary for the addressee to know and understand that a given response is to be expected. So, in using *si* in such circumstances, the speaker is signalling the fact that the called-for response is the normal, expectable thing.

3) Utterances urging response to some noteworthy sensual stimulus. For example, a speaker urges the addressee to look at a pretty girl, or to listen to an odd sound, or to feel the texture of a luxurious piece of cloth. In situations of this sort, the addressee knows what is expectable as soon as the speaker has expressed himself. Why it is expectable he will learn as soon as he has done what he is being urged to do. In any case, the speaker need make no explanation (unless he wishes) to justify the action he is calling for. The situation speaks for itself.

#### 2.2.3.1.2. 'Expectability' that Requires Explanation or Justification

In contrast to the above situations, there are other cases where the expectability of a given response is not self-evident from the command or suggestion as it stands. In such cases the speaker must provide some explanation or suggest some consideration that clarifies why the given action is called for in this particular instance.

Explanations and clarifications of this sort are many and varied. They may comprise offers of reward (example 4), or comments about some desirable consequence of the action in question. Or they may take the form of threats or of warnings about undesirable consequences.

They may also point to some new fact or event in the immediate environment (such as the time of day, the weather, or the ringing of the doorbell) that may not have been noticed by the addressee, but that gives rise to a call for the action in question. Then again they may point back to old information as an inducement to the action. For example, the addressee has been offered a gift (so he should accept it; see item 6); or he has just indicated that he wants to go to the party (so he should go); or he has just noticed how rickety a chair is (so he shouldn't sit in it; item 7).

Commands and suggestions of this latter type (i.e. those referring back to old information) differ from the others in that the reason for the called-for action, being old information to the addressee, need not be explicitly stated as a part of the action-inducement utterance. But that reason will usually be acknowledged or signalled as a part of the command by means of the addition of the preverbal particle /kâ .../ *'then, well then ...'*. (Compare examples 6 and 7.) The mandatory presence of /kâ/ is here taken, then, to mark such utterances as falling into the category of commands or suggestions of the type whose expectability is not self-evident but must be explained or justified in some way.

Among the most common of the situations calling for explanations or clarifications of the kind mentioned above are those in which the addressee seems to be unaware of or heedless of some crucial fact or consideration that the speaker feels should govern his actions. For example, the addressee doesn't seem to realise how good the prices are at a particular store; or he evidently didn't hear the doorbell; or he is unaware of the time; or he is not sufficiently heedful of the significance of the fact that a gift is being offered to him. He therefore needs these considerations brought to his attention if he is to be urged (using the particle *si*) to shop at that store, to answer that doorbell, to hurry and get dressed for that scheduled event, or to accept that gift offered to him. Once these considerations are pointed out in some way, the action in question then becomes the obvious, expectable thing to do, at least in the speaker's eyes.

A particular subgroup of situations of the above sort consists of those situations which give rise to preventive suggestions or commands. These are situations in which the addressee seems to be about to do



something uncalled for, evidently unaware of some crucial consideration that would otherwise prevent him from acting as intended. For example, a child is about to touch the stove, unaware that he might get burned; or a friend is about to shop at a particular store without realising how dishonest the establishment is; or a guest seems to feel obligated to sit in a rickety chair ignoring or suppressing his own doubts about its serviceability. So the speaker, using the particle *si*, urges the addressee not to perform the given action, and at the same time provides the explanation or points to the consideration that makes the addressee's response expectable.

Preventive negative commands of this kind thus contrast with corrective negatives (in which the speaker tells the addressee not to do something he is already doing) in that the former require some justification for the prohibition (if not an explicit reason, then at least the form /kǎ .../ 'well then ...') whereas the latter do not. They also contrast with preventive commands using the particle /ná/, where the speaker is simply expressing his wishes or demands, and therefore needs append no explanation or justification.

#### 2.2.3.2. 'Expectability' in Speaker-response Utterances

Up to this point, we have been dealing with the matter of expectable response as it relates to action-inducement utterances, or utterances that call for an expectable or obvious response on the part of the addressee. Let us now look at responses on the part of the speaker. Here we find two new aspects to the problem of obviousness or expectability. First is the fact that the expectability need not always be clear (or made clear) to the addressee, for it is the speaker's own response that is being judged expectable, not that of the addressee. Thus in example 13 above, the addressee has no idea whether the third party is coming or not. But the speaker knows; and when the speaker answers /maa sǐ/ 'Sure he's coming', he conveys the idea that he has what he considers sufficient reason for his response; but he need not explain the reasons for his assurance to the addressee. He can explain the basis for his assurance if he wishes, but even if he does not, he can still signal his assurance by the use of /sǐ/; and the addressee will know that the response is based on what the speaker feels to be good grounds.

The second aspect to the problem of expectability or obviousness is the fact that in many instances, including the above example, the speaker is basically conveying the idea that he is sure, or has reason to believe, that something is the case. In other words, /sǐ/ here carries the idea,

not strictly of expectability or even obviousness, but of assurance, certainty, and behind that of the presence of a reason for that assurance or certainty.

Let us now look at the various types of speaker-response occurrences of /sî/.

#### 2.2.3.2.1. Answers to Questions

As we have seen (examples 13-17), /sî/ may occur with answers to questions. But the characteristics of usage will vary depending upon whether those questions are yes-no questions (i.e. questions which call for a yes or no answer) or content questions (i.e. questions asking who, what, when, how many, why, etc.).

The chief difference between yes-no questions and content questions, with respect to the use of /sî/, is the fact that the particle never occurs in answers to content questions unless the information in the answer is something that is known or ought to be known to the addressee; and then /sî/ always occurs as a part of the expression /kô ... na sî/ 'well ...'. (See examples 16-17.) But in the case of yes-no questions, /sî/ may occur not only in contexts of this sort, but also in situations where the facts of the matter are known only to the speaker; so /sî/ may occur either with or without /kô ... na/, depending on the context. (See examples 13-15.) I must confess that I do not know why the difference between yes-no questions and content questions should give rise to this difference in the use of /sî/; but I suspect that the problem has something to do with some unique semantic characteristic inherent in content questions that so far has eluded me.

There is also a particular limitation upon the use of /sî/ in answers to yes-no questions of the type where the questioner has some expectations about the response. This includes yes-no questions signalled by /lăə/ or /chây máy/; for example /khăw<sub>1</sub> rúu<sub>2</sub> lăə<sub>3</sub>/ 'He<sub>1</sub> knows<sub>2</sub>, huh?<sub>3</sub>', or /khăw máy rúu lăə/ 'Doesn't he know?', or /khăw rúu chây máy/ 'He knows, doesn't he?'. In negative answers to questions of this sort, the speaker will ordinarily use /sî/ only if he at the same time signals in some way the reason for his response. This he may do by actually stating the reason, or (if he feels the reason ought to have been already clear to the addressee) by means of the expression /kô ... na sî/ 'well ...'. If on the other hand the speaker makes no reference to the reason for his response, he will ordinarily use the particle /rôk/ (signalling a contradictory or occasionally a confirming negative response), but never /sî/.

## 2.2.3.2.2. Responses to Question-raising Statements

As mentioned earlier, these are responses not to questions asked by the addressee, but to statements which for the speaker raise a question of fact or perception; and these statements then evoke some kind of confirming, contradictory, or explanatory response from the speaker.

The possibilities of and restrictions upon the occurrence of /sî/ in such responses may be illustrated by the following examples of possible and impossible responses to item 30 below. Items a, c, d, e, g, h, represent possible responses to the sentence; but items b and f, marked by an asterisk, are unacceptable:

- (30) /phôm<sub>1</sub> khít<sub>2</sub> wâa<sub>3</sub> 'aacaan<sub>4</sub> ca<sub>5</sub> sùu<sub>6</sub> rôt<sub>7</sub>yon-/ 'I<sub>1</sub> think<sub>2</sub> that<sub>3</sub> the professor<sub>4</sub> will<sub>5</sub> buy<sub>6</sub> a car<sub>7</sub>.'
- a. /sùu sî/ 'He certainly will.'
- b. \*/mây sùu sî/ 'No he won't.'
- c. /mây sùu sî, khâw<sub>1</sub> mây<sub>2</sub> mii<sub>3</sub> nēn<sub>4</sub> pho<sub>5</sub>/ 'No he won't, he<sub>1</sub> doesn't<sub>2</sub> have<sub>3</sub> enough<sub>5</sub> money<sub>4</sub>.'
- d. /khuan<sub>1</sub> tōn<sub>2</sub> pen<sub>3</sub> mōtēsāy<sub>4</sub> sî/ 'It'll surely<sub>1</sub> have<sub>2</sub> to be<sub>3</sub> a motor cycle<sub>4</sub>.'
- e. /phîi<sub>1</sub> 'aacaan<sub>2</sub> sî ca<sub>3</sub> sùu<sub>4</sub>/ 'It's the professor's<sub>2</sub> brother<sub>1</sub> that's going to<sub>3</sub> buy<sub>4</sub> (one).'
- f. \*/'aacaan<sub>1</sub> sî ca<sub>2</sub> sùu<sub>3</sub>/ 'It's the professor<sub>1</sub> that's going to<sub>2</sub> buy<sub>3</sub> (one).'
- g. /khon<sub>1</sub> yān<sub>2</sub> 'aacaan<sub>3</sub> sî tōn<sub>4</sub> sùu<sub>5</sub> nēn<sub>6</sub>/ 'A person<sub>1</sub> like<sub>2</sub> him<sub>3</sub> has got to<sub>4</sub> buy<sub>5</sub> (one) for sure<sub>6</sub>.'
- h. /thāa<sub>1</sub> mii<sub>2</sub> thūrá<sub>3</sub> māk<sub>4</sub> yān<sub>5</sub> nān<sub>6</sub> kō tōn<sub>7</sub> mii<sub>8</sub> rōt<sub>9</sub> sūan tua<sub>10</sub> sî/ 'If<sub>1</sub> (you) have<sub>2</sub> a lot<sub>4</sub> of business<sub>3</sub> like<sub>5</sub> that<sub>6</sub>, (you) have to<sub>7</sub> have<sub>8</sub> (your) own<sub>10</sub> car<sub>9</sub>.'

The first thing to be noted from the above examples is the fact that in responses to question-raising statements, as in answers to questions, the reason for the expectable response need not be clear to the addressee; and in most instances the speaker will not make it clear. In other words, when the speaker uses /sî/ in such responses, the addressee knows that the former has good reason for his response, but he need not know, and perhaps will not even be told what those reasons are. So, here again (as with answers to questions), the particle /sî/ may convey

assurance, certainty on the part of the speaker, based on private reasons which seem to him to be adequate.

Also to be noted is the fact that responses to question-raising statements include some responses that address themselves to the truth or falsehood of the addressee's utterance (see 18, and 30 a,c), and others that are concerned with the question of who, what, why, etc. (see 20, 21, 22, 30d, e, g, h). Clearly the former are rather like answers to yes-no questions in this respect, and the latter are comparable to answers to content questions. The who-what type of response here, however, does not (as in the case of answers to questions) require the occurrence of /k̂ ... na sî/ 'well ...'. In fact either type of response can occur with /sî/ by itself; and either can occur with /k̂ ... na sî/ in situations where the speaker feels the addressee should or could have known the facts of the matter.

Still another matter to be noted is this: that responses of the type under consideration may be contradictory, confirmatory, or simply explanatory. Contradictory responses (for example 18, 20, 30c, d, e, and potentially h) are those in which the speaker feels called upon to contradict or correct something the addressee has said or implied. Such occurrences may occur freely with /sî/, except that negative responses of the yes-no type, whether contradictory or not, must ordinarily be accompanied by some kind of explanation or signal pointing to the reason for the negative response. If such reference to the reason is missing, then the speaker will usually use /r̂k/ (contradictory or confirming negative) rather than /sî/. Thus example 30c is permissible, but not 30b. Example 30b would, however, be an acceptable response if it were marked by the particle /r̂k/ rather than /sî/.

Confirmatory responses (e.g. 30a, g, h, and potentially 21, 22) are those which agree in essence with what the addressee has just said. Positive confirmatory responses of the yes-no type of response can occur freely, no elaboration or explanation being necessary. However, a confirmatory who-what type of response can only occur if the speaker somehow enlarges upon what the addressee has just said. This enlargement can take the form of a generalisation (21, 22, 30g), or it can provide an explanation of some sort (19, 30h). But confirmatory responses of the type exemplified in 30f cannot occur, presumably because they involve an inappropriate topicalisation of the discourse subject.

Explanatory responses explain why or how some fact or state of affairs alluded to by the addressee is or should be as noted. Such responses are usually marked by the occurrence of the preverb particle /k̂/ (see 19, 30h).

One last thing to be noted is the fact that in some responses the particle /sɨ/ occurs in the middle of the sentence (see 20, 21, 22, 30e, g). All such cases turn out to be what I call topic-focus sentences, that is who-what sentences in which the subject or topic of the sentence forms the focus of the sentence predication. So they all convey the idea that it's the subject of the sentence about which something can properly be affirmed; i.e. it's that person that's pretty (20); it's this kind of rain that causes floods (21), it's being a child that's pleasant (22), etc. We can therefore make the general statement that /sɨ/ always immediately follows the focal or central predication of the sentence. In most types of sentences, the particle therefore appears at the end of the sentence, but not so in the case of subject-centred predications of the type exemplified above. Strictly speaking, then, /sɨ/ is not a sentence-final particle at all, but a predication marker of some sort.<sup>4</sup>

#### 2.2.3.2.3. Inferential Statements

These comprise utterances in which the speaker responds to certain facts or clues that have come to his attention, by drawing some inference or stating some conclusion that he arrives at on the basis of those clues. For example, the addressee starts elaborating his plans for a party at the speaker's house, and this provides the latter with the clue that leads to the conclusion that the addressee must be expecting the speaker to participate - a conclusion that had not previously been stated explicitly; so the speaker makes the appropriate inference (see 23). Or again, the addressee's problems with 5:00 p.m. traffic lead the speaker to infer that there must be a typical rush-hour traffic jam (see 24). Or still again, the speaker sees the addressee coming in with a dripping umbrella and raincoat, and infers that it must be raining outside (see 25).

Note, however, that these utterances must be inferences, not first-hand observations of fact. If the addressee in the first example above has specifically stated his expectations, or if the speaker is actually observing the traffic jam, he will not use the particle /sɨ/ (unless he is noting the information for the first time, and then he may respond as in the statements noting new information to be discussed below). Also, for some speakers, the inference or conclusion to be drawn must be a fairly clear one. If it is somewhat doubtful or tenuous, such speakers would ordinarily use the mid-tone form /si/ (see later discussion).

#### 2.2.3.2.4. Statements Noting New Information

As stated above, these comprise responses to some new development or fact that has just come to the attention of the speaker, or that is being brought to the attention of the addressee for the first time. And such statements always have the form /lɛɛw/ 'now, already' as a part of the predication. The idea of reasonable, necessary, or expectable response is a little more obscure in utterances of this type, but it is, I believe, nevertheless present. The implication of /sɪ/ here, is that the new fact brought to light must now - reasonably, expectably, necessarily - be at least noted and also (where appropriate) adjusted to, reckoned with. So, behind the utterance there is, as it were, a veiled command or suggestion to the speaker or addressee or both to see, hear, take note, consider, adjust, or whatever. This is the expectable, necessary, appropriate thing to do.

### 3. FORMS AND MEANINGS OF OTHER VARIANTS OF SI

So far we have been considering only the phonetic form /sɪ/ and the variety of contexts in which it occurs and the basic meaning which it has in all those contexts. Now we are in a position to consider other phonological forms of this same particle: /si/ or /sii/, /sɪi/, /sɪ/, and /sfi/. These forms, along with /sɪ/ are all variants of the particle si; and each of these variants retains the basic meaning of the particle, but each also has a further semantic value that distinguishes it from all the rest.

#### 3.1. SPECIAL PHONETIC CHARACTERISTICS OF VARIANTS

The phonetic values of the above-mentioned variants, as it turns out, differ in a number of ways from the values of other comparable non-particle forms in the language. For one thing, the vowels of the various forms of si are often pronounced lower and more lax than other non-particle forms ending in /-i/ or /-ii/. Also, the short vowel forms /si/, /sɪ/, and /sɪ/ never under any circumstances are pronounced with a terminal glottal stop, whereas other comparable forms in the language usually are when they occur in terminal or stressed position. Then the falling-tone forms /sɪ/ and /sɪi/ (unless particularly stressed or emphasised) drop from the mid-tone level or even lower, whereas other falling-tone forms usually drop from the high-tone level or even higher. The form /sɪ/, in particular, when unstressed, can drop from the mid level to a little below mid, or to low, or to any point between; or it can drop from lower mid or even low. Incidentally, other discourse

particles (though not all of them) share many of the unusual characteristics described above.

All these peculiarities of *si*, added to the elusiveness of semantic distinctions and the variability of vowel-length under conditions of stress variation (a common enough phenomenon in the language), give rise to considerable difficulty in differentiating the variants of the particle or in determining which variant is present in a given utterance. In this regard, the distinction between /sǐ/, /si/, and a hypothetical /sì/ has presented the most difficulty. As it happens, not all speakers distinguish consistently between /sǐ/ and /si/ (unless the former occurs particularly stressed or emphasised); and, so far as I can tell, /sǐ/ and /sì/ never clearly contrast; so the latter could probably be considered a freely varying allomorph of /sǐ/.

### 3.2. EXAMPLES OF USAGE OF CONTRASTING FORMS

We are left, then, with the forms /sǐ/, /si/ or /sii/,<sup>5</sup> /sǐi/, /sf/, and /sǐi/ as variants which are distinguished from each other both in form and meaning. The semantic similarities and differences between them may be illustrated by showing what happens to the sentence /pǎet, pratuu<sub>2</sub>/ 'Open<sub>1</sub> the door<sub>2</sub>.' when it is accompanied by each of the variants:

- (31) /pǎet pratuu sǐ/ 'Open the door.' [It's time for the addressee to open the store door.]
- (32) /... si/ or /... sii/ 'Hey, how about opening the door!' [The addressee should be opening the door, but he is hanging back or woolgathering.]
- (33) /... sǐi/ 'Come on, do open it!' [The addressee is refusing to open the door; or he has ignored one or more previous requests.]
- (34) /... sǐ/ 'Open the door, would you!' [The speaker wants to be able to look inside the room or closet, or he wants to put something away, and he needs the addressee to open the door for him; but note that the Thai utterance is not strictly a question.]
- (35) /... sǐi/ 'PLE-EASE open the door!' [The speaker is a child who desperately wants to get into the bathroom, and his older brother is teasing him or refusing to let him in. Some speakers would simply use /sǐi/ here, with raised pitch to indicate emphasis, insistence, urgency.]

### 3.3. CONTRASTING MEANINGS OF VARIANTS

Concerning the similarities between the above sentences, it is sufficient for the present to say that each conveys the idea that the opening of the door is the obvious, reasonable, expectable thing for the addressee to do under the circumstances. In other words, *si* in all its phonologically variant forms, as exemplified above, still retains this meaning of obvious or expectable response. But what semantic differences are signalled by these variations in form? This is the question to which I shall now address myself.

#### 3.3.1. Contrasting Meanings as Handled in the Literature

Unfortunately, most sources in the literature provide rather little help at this point. Thus several authorities simply list two or three phonological variants without specifying what the differences in pronunciation mean - which conveys the impression, perhaps unintentionally, that the different forms vary freely with no particular significance to be assigned to each variant. See McFarland (1954), Thai-Thai Dictionary (1976), Manitcharoen (1964), Haas (1964), Noss (1964), Bhamoraput (1972).

Henderson (1949), on the other hand, attempts to explain the phonological forms of all the sentence-final particles by describing various prosodic features of length and pitch, and listing possible combinations of these features, and then assigning general meanings to each combination. For example, she suggests (p.207) that a short falling-pitch combination conveys "assertion, or assent, or command", while a short high pitch conveys "interrogation, invitation", and so forth. But as it turns out, almost all her generalisations have exceptions, and besides they are too general to provide much help for understanding the variations in form and meaning of particular particles.

Chuenkongchoo (1956) carries the matter a little further, giving examples of utterances where each variant of each particle occurs, and going into a little more detail than Henderson about general meanings of the various prosodic combinations. But again his generalisations provide only limited help for understanding the varying forms and meanings of particular particles. One comment of his, however (1956:70), does seem to be at least partially applicable to the forms /sii/ and /síi/: "Length," he says, "is often used to add 'intensity' or extra weight to utterances in which in other contexts a short particle might be used. Situations involving 'insistence' or 'exasperation' frequently call for complexes in which length is a feature."

Rudaravanija (1965), like Henderson and Chuenkongchoo, proposes generalised meanings for different phonological characteristics of final



particles. But she carries the matter further by suggesting meanings for varying pitches of a few individual particles. For example (p.95), she informs us that *na* with rising terminal contour has a 'mild emphatic' meaning; and with falling contour it is 'strong emphatic'. However, her semantic generalisations about final particles focus upon the feature of pitch or final contour, and not upon other features such as length or terminal glottal closure. And, unfortunately for our purposes, she omits *si* from her discussion; so we are left without the benefit of her judgement in this particular case.

The clearest and most specific help, in my opinion, comes to us from Brown (1969:3.20) in his definitions of /sɪ/ and /sɪ̃/. According to him, /sɪ/ is "a particle used to request an action when the result of the action, not the action itself, is the point of the request"; and /sɪ̃/ is "a particle used to request or urge an action when the action itself is the point of the request." And that's about all the really helpful information I have been able to find in the literature.

### 3.3.2. Examples and Suggested Meanings of Each Variant

Let me now turn to some further examples of each of the variant forms, and then some definitions and explanations. As before, the following examples, for the most part, reflect the speech of Ms Kiranand. No doubt many speakers will differ from her at one point or another. For example, one speaker, Ms Bandhamedha, with whom I have worked extensively, makes no distinction as below between /sɪ̃/ and /sɪ/ or /sɪi/; and other speakers make use of the form /sɪ̃i/, while Ms Kiranand does not. However, I believe Ms Kiranand's usage is not particularly idiosyncratic, and it will serve as a convenient basis for presenting the data.

#### 3.3.2.1. The Form /sɪ/ or /sɪi/<sup>6</sup>

- (36) /khǎn<sub>1</sub> hâi<sub>2</sub> dii<sub>2</sub> sii, lɛ̃w<sub>3</sub> ca<sub>4</sub> dâay<sub>5</sub> raanwan<sub>6</sub>/ 'Write<sub>1</sub> nicely<sub>2</sub> now, and<sub>3</sub> (you)'ll<sub>4</sub> get<sub>5</sub> a reward<sub>6</sub>.' (cf. example 4, above.)  
[The mother holds out a reward to her child as an inducement for writing a nice letter, and then she withdraws it as the child reaches for it, thus conveying the idea that the reward will not be his until the letter is written to her satisfaction.]
- (37) /khǎw<sub>1</sub> hâi<sub>2</sub> kô<sub>3</sub> ʔaw<sub>4</sub> sii/ 'He<sub>1</sub>'s giving<sub>2</sub> it to you, so<sub>3</sub> take<sub>4</sub> it.' (cf. example 6.) [The speaker is baffled, and perhaps a little annoyed that the addressee is hesitating. Or: The addressee has asked the speaker what to do, and the latter doesn't want to be bothered with the problem.]

- (38) /khâu maa sii/ 'Come in.' (cf. example 8.) [The speaker is not really too eager to have the addressee come in. Or: The speaker knows that the addressee has come to see someone else, so the speaker is not involved in the business or pleasure for which the addressee has come.]
- (39) S1 /khun<sub>1</sub> ca<sub>2</sub> mây<sub>3</sub> klàp bân<sub>4</sub> lă<sub>5</sub>/ S2 /klàp sii/ 'Aren't<sub>3</sub> you<sub>1</sub> going to<sub>2</sub> go home<sub>4</sub>? 'Sure I am.' (cf. example 15.) [The second speaker feels the first shouldn't have had to ask. Or: The second speaker doesn't want to be bothered with the problem.]
- (40) /khun<sub>1</sub> khít wâa<sub>2</sub> chán<sub>3</sub> tòk lo<sub>4</sub> sii/ 'You<sub>1</sub> seem to think<sub>2</sub> I<sub>3</sub> agree<sub>4</sub>.' (cf. example 23.) [The speaker gathers from the addressee's manner or behaviour that the latter expects him to agree to having a party at his house. The inference drawn by the speaker here is more doubtful or tentative than that in comparable example for /sî/, item 20. Thus inferential statements with /sii/ have the general flavour of English utterances accompanied by phrases such as 'I guess', 'I suppose', 'it seems as if'. They also have a slight hint of questioning about them, though not to the point of requiring any response from the addressee.]

The meaning conveyed by /sî/ or /sii/, as in the above examples, is that of ininvolvement, indifference, emotional neutrality. And this uninvolvement may be simple and straightforward, or it may be an assumed indifference that both masks and expresses anything from mild to strong coldness, withdrawal, rejection, hostility. The simple kind of uninvolvement is exemplified in one of the possible situations where example 38 might occur. Here the speaker is in fact not involved (and is not expected to be) in the invitation extended to the guest. It is also exemplified in utterances such as 40, where the speaker is making a tentative inference on the basis of clues he thinks he has picked up. In other words, he is not jumping to a definite conclusion - as he would be if he were using the form /sî/. Thus /sii/ renders the inference much more indefinite and non-committal.

The other examples above illustrate the use of /sii/ to express the more emotionally-loaded kind of non-involvement. Thus, in example 36, the mother is in effect withdrawing emotionally from her child, and she expresses this fact both by the use of /sii/, and by her withdrawal of the promised reward from the child's outreached hand. And in the other examples the speaker is expressing a non-involvement that both conceals and reveals his impatience and hostility: why doesn't the addressee open the door as expected (example 32), or take the gift that's being

offered (example 37)? Or why does the speaker have to be bothered with the question (example 39)? But note that the impatience or hostility is that of emotional coldness, withdrawal, or uninvolvedness, not that of emotional heat or aggression. The latter would be expressed by /sî/, with falling tone, and raised above normal pitch.

If we consider the situations in which it is possible to use the form /si/ or /sii/, and thus convey non-involvement or emotional neutrality, we find that most of the situations that allow /sî/ also allow these mid-tone forms. In other words, in most situations exemplified and discussed above (see examples 1-29), there can be a formal and semantic contrast between /sî/ on the one hand, and /si/ or /sii/ on the other. Exceptions are as follows: Negative commands (as in 2) and also topic-focus statements (see 20, 21, 22, 30e, g) only occur with /sî/, never with /sii/. Also I have not been able to elicit a /sii/ counterpart for example 18. On the other hand, statements noting new information (see 26-29) may occur with either /si/ or /sii/, with no difference in meaning between the two. Similarly, all utterances accompanied by /kô ... na sî/ (see 14, 16, 17) may occur with /kô ... nâ si/, again with no differentiation in meaning between the two. I cannot adequately account for the exceptions listed above.

### 3.3.2.2. The Form /sîi/

- (41) /yàa<sub>1</sub> khàp<sub>2</sub> rew<sub>3</sub> sîi/ 'Please don't<sub>1</sub> drive<sub>2</sub> so fast<sub>3</sub>.' 'Do slow down, for goodness sake!' (cf. item 2.)
- (42) /yìp<sub>1</sub> dɪnsɔ̌<sub>2</sub> hây nòy<sub>3</sub> sîi/ 'Aw, come on, please<sub>4</sub> reach<sub>1</sub> (me) the pencil<sub>2</sub>.' (cf. item 10.) [The speaker has asked for the pencil before, but the addressee was too lazy to get up, or he is teasing the speaker.]
- (43) S1 /kháw<sub>1</sub> ca<sub>2</sub> maa<sub>3</sub> cɪŋ cɪŋ<sub>4</sub> lǎə<sub>5</sub>/ S2 /maa<sub>3</sub> sîi/ 'Will<sub>2</sub> he<sub>1</sub> really<sub>4</sub> come<sub>3</sub>?<sub>5</sub>' 'Su-ure he will.' 'Why certainly he will.'
- (44) S1 /chán<sub>1</sub> wâa<sub>2</sub> wanní<sub>3</sub> fǎn<sub>4</sub> thâa ca<sub>5</sub> mây<sub>6</sub> tòk<sub>4</sub>/ S2 /tòk sîi/ 'I<sub>1</sub> think<sub>2</sub> it probably<sub>5</sub> won't<sub>6</sub> rain<sub>4</sub> today<sub>3</sub>.' 'Aw come on now, sure it will.' [The second speaker has previously given his reasons for thinking it will rain, but the first speaker evidently still won't see reason.]

The meaning conveyed by /sîi/ is the idea of persuasion, exerting pressure. Usually this form will occur in situations where the addressee has been unnecessarily slow in complying with the speaker's expectations, either in action or belief. It would therefore be very natural to use /sîi/ in an interchange such as the following:

- (45) S1 /yìp<sub>1</sub> dɪnsǎ<sub>2</sub> hây<sub>3</sub> nòy<sub>4</sub> sǐ/ 'Hand me the pencil would you.'  
 (reach<sub>1</sub> pencil<sub>2</sub> for (me)<sub>3</sub> a little<sub>4</sub>) S2 (ignores the request)

S1 /yìp !sǐ/ 'Hand it to me!' (The exclamation point here indicates raised pitch.) S2 /chán<sub>1</sub> khiikiat<sub>2</sub> yìp<sub>3</sub>/ 'I<sub>1</sub> am (too) lazy<sub>2</sub> to bother.'

S1 /yìp hây nòy sǐi/ 'Come on now, do hand it to me!'

The form /sǐi/ occurs only in action-inducement utterances (positive or negative), and in responses to questions or statements. It does not usually, however, occur with topic focus statements (see 20-22, 30e, g), or in /kô ... na .../ utterances (see 14, 16, 17).

### 3.3.2.3. The Form /sǐ/

- (46) /yìp<sub>1</sub> dɪnsǎ<sub>2</sub> hây nòy sǐ/ 'Reach<sub>1</sub> me the pencil<sub>2</sub>, would you.'  
 [The speaker needs the pencil and avails himself of the addressee's help to meet his needs.]
- (47) /fan<sub>1</sub> sǐ/ 'Listen<sub>1</sub>!' 'Listen, would you.' [The speaker can't hear someone who is talking, and he wants the addressee to listen and see if he can catch what's being said. Or: The speaker can't identify some sound and he wants the addressee to help him out.]
- (48) /ʔâa<sub>1</sub> pàak<sub>2</sub> sǐ/ 'Open<sub>1</sub> your mouth<sub>2</sub>.' [A dentist is speaking to his patient. Or: A mother wants to see what her child has in his mouth. Or: A Thai language teacher wants to test the vocabulary command of a non-native pupil.]

The form /sǐ/, as in the above examples, conveys the idea that the action called for from the addressee is needed or desired by the speaker for some purpose of his own. This need or desire may be a matter of personal comfort, curiosity, or even whim; or it may involve something the speaker needs to have done so that he can in turn do something else that he (or the addressee, or someone else) wishes to have done. This form is used only with action-inducement utterances, and then only in the positive. That is, it never occurs with /yâa .../ 'do not ...'.

### 3.3.2.4. The Form /sǐi/

- (49) /yìp<sub>1</sub> hây<sub>2</sub> nòy<sub>3</sub> sǐi/ 'Reach (it)<sub>1</sub> for (me)<sub>2</sub>, ple-e-ase<sub>3</sub>!' [A child is getting very impatient and insistent to an older sibling who is ignoring him or teasing him by not reaching for something the child needs and can't reach for himself.]

The form /s<sup>h</sup>i/ conveys both the idea of the speaker's need or wish, and also the idea of persuasion, pressure, insistence, and sometimes even urgency. Like /s<sup>h</sup>/ it is used only in positive action-inducement utterances. Furthermore, it is chiefly used by children, and to a lesser extent by women. I should point out, however, that not all speakers accept this form. In fact it is the one variant that does not occur in Ms Kiranand's speech. Possibly all occurrences of /s<sup>h</sup>i/ should be interpreted as occurrences of /sii/ which have been raised extra high. (For discussion of the phenomenon of raising, see section 4, below.)

This brings us to the end of our discussion of what may be considered the basic variants of si. There now remain two further types of form and/or meaning variation that require our consideration: the phenomena of raising and of neutralisation.

#### 4. THE PHENOMENON OF RAISING

Raising may be defined here as the process in which the pitch of a falling-tone particle (in this case /s<sup>h</sup>i/ or /sii/) is elevated above the normal level in order to convey emphasis or increased emotional intensity. Thus the particle /s<sup>h</sup>i/ in the utterance /p<sup>h</sup>æt pratuu s<sup>h</sup>i/ '*Open the door.*' might undergo raising if the speaker were particularly annoyed, or if he had to repeat the suggestion or command a second time. Such raising would then be indicated in the transcription by means of an exclamation symbol immediately preceding the raised form: /p<sup>h</sup>æt pratuu !s<sup>h</sup>i/. (See also examples 2, 18, 45, above.)

Raising, as described above, should be distinguished from two other types of raising that occur in the language. In one of these a syllable of any tone is changed from its normal pitch to an extra high and slightly rising pitch, thus expressing a particular kind of emphasis (see Haas 1964:xii-xiii). In the other type the pitch of the whole sentence is raised above the normal level (see Haas 1964:xiii). By way of contrast, the type of raising that here concerns us has its effect exclusively upon falling-tone particles or particle variants. Such particles, when raised, are pitched above their normal range, but they still retain their falling contour.

Now, as has been noted above (3.1.), the normal, unraised pitch of /s<sup>h</sup>i/ and /sii/ (and also, incidentally, of other falling-tone particles) is lower than that of falling-tone non-particle forms in the language. Thus, when unraised, these particle forms will start off from a point at or below the normal mid-tone level, and then drop on down from there. So they can drop from mid to lower mid or to low, or from lower mid to

low, or even from low to a little lower still. But, when raised, these forms fall from a starting point above the mid-tone level. And the raising, furthermore, is variable. That is, the pitch may be raised just a little, or it can be raised a great deal; but the higher the raising, the greater the degree of emphasis or intensity conveyed. There is, however, a clear dividing line between raised and unraised forms. Thus if the pitch falls from the mid-tone level or lower, the form is unraised; but if it falls from a starting point above the mid-tone level, then it is raised; and such raising therefore conveys the concomitant semantic value accordingly.

Any use of /sɪ/ or /sɪi/ which is clearly assertive, contradictory, rebuking, etc. will be raised. Thus the particle /sɪ/ in examples 2 and 18, above, is of necessity raised. This is so in the former case because the sentence in question is a flat, negative command; and, like all negative /sɪ/ commands, it necessarily implies rebuke for some undesirable action. Then in the latter example (18), the sentence is a flat contradiction. So both are examples of the kind of assertiveness that calls for raising above the normal pitch of the particle variant /sɪ/.

Of the various types of occurrence of /sɪ/ exemplified in 2.1. above, raising may occur with action-inducement utterances (cf. examples 1-12), and with responses to questions (cf. 13-17), and to question-raising statements (cf. 18-22). But raising cannot occur with inferential statements (see 23-25) or with statements noting new information (26-29).

Incidentally, in the case of topic-focus utterances (see examples 20-22), /sɪ/ can be raised only if the sentence in question constitutes a contradictory or assertive statement insisting that 'subject A' (not 'subject B') is the one of whom some predication may properly be made. In view of this requirement, examples 21 and 22 cannot, as they now stand, be raised in any of the given illustrative contexts; but contexts could be framed such that raising might indeed occur. Thus, for example, the speaker in utterance 21 might have been arguing with the addressee about what kind of rain causes floods; and if he is annoyed with the other's obtuseness concerning the obvious danger of this kind of rain (as opposed to some other kind the addressee insists on stressing), he then can use a raised /sɪ/ to make his point.

An interesting demonstration of the importance of the distinction between raised and unraised forms appears in the speech of one of my language assistants. Ordinarily this speaker makes no distinction between /sɪ/ (straightforward meaning) and /sɪi/ expressing non-involvement), and she perceives all occurrences of these in her own

speech as having mid tone. In other words, most of the examples listed in items 1-29 and 36-40 are so perceived. And this is true regardless of the fact that in her own speech these occurrences may be variously pronounced with mid pitch, or low, or mid falling to lower-mid or to low, or lower-mid falling to low. But if she pronounces the particle in raised fashion, that is with a pitch starting above the mid level and then dropping on down, she immediately identifies it as having falling tone. And such forms then convey emphasis or intensified emotion - with the expectable corollary that none of the non-involvement utterances (see 36-40) can occur with falling tone. In other words, her mid-tone /si/ (often pronounced with falling pitch) corresponds to other speakers' falling-tone /sî/ and to their /si/ or /sii/; and her falling tone /sî/ corresponds to their raised falling tone /!sî/. Furthermore, her /si/ is perceived as having mid tone even when it drops, provided it doesn't drop from a point higher than mid tone. If it does drop from a higher point (i.e. the point which divides raised from non-raised forms for other speakers), it will be perceived as having falling tone. Strangest of all is the fact that other falling-tone particles, such as /khâ/ (female deference) and /nâ/ (old information), are perceived as having falling tone despite the fact they may be pronounced in ways exactly parallel to her non-raised pronunciation of /sî/, or /si/. My guess is that these perceptions are a product of the Thai writing system complicated by some kind of interference from intonational phenomena. In any case, these special perceptions are not a reflection of any inability on her part to hear the phonetic facts, for she recognises these when they are pointed out to her. It seems to be tied in with intuitive perceptions of some kind.

##### 5. THE PHENOMENON OF NEUTRALISATION

The second phenomenon (besides raising) that requires consideration is that of neutralisation. This term refers to a process in which the potential variability of si, both with respect to form and meaning, is neutralised or blocked, leaving /si/ as the only permissible alternative. Such neutralisation takes place whenever si is immediately followed in the sentence by another particle - usually one of the status-intimacy particles such as /cá/, /khá/, etc. And for most speakers oddly enough, it is always the question form of these status-intimacy particles that occurs, never the statement form.<sup>8</sup>

To explain further, if the different sentences cited in examples 31-35 were to be altered by the addition of /khá/, only the form /si/ would be permissible in each case. Furthermore, all semantic differentiation

would be lost, so that the resulting sentences would convey a rather neutral sense of what is expectable. Then the added /khá/ would further convey the sense of politeness or deference.

As already mentioned, when *si* is followed by another particle, thus producing neutralisation, the following particle will usually be one of the status-intimacy forms such as /cá/, /khá/, etc.; but the particle *na* can also occur:

- (50) /ní<sub>1</sub> khon<sub>2</sub> 'íik<sub>3</sub> láay<sub>4</sub> wan<sub>5</sub> si ná<sub>6</sub> kwàa ca<sub>7</sub> sèt<sub>8</sub>/ 'But<sub>1</sub> it'll  
surely<sub>2</sub> be several<sub>4</sub> more<sub>3</sub> days<sub>5</sub> before<sub>7</sub> it's finished<sub>8</sub>, won't it<sub>6</sub>?'

As it happens, most other particles which end in a short vowel will undergo partial or complete neutralisation under similar circumstances. This means that, ordinarily, only one particle in a series - usually the last one - appears in its developed or unneutralised form. However, the forms /ná/ (old information) and /ní/ or /níi/ (new information) seem to be exempt from the necessity of neutralisation, as may be seen in the following example:

- (51) /kháw<sub>1</sub> klàp baan<sub>2</sub> pay léew<sub>3</sub> ní<sub>4</sub> si/ 'But<sub>4</sub> he<sub>1</sub>'s already<sub>3</sub> gone  
home<sub>2</sub>.'

## 6. POSSIBILITIES AND RESTRICTIONS WITH RESPECT TO USAGE

And now, before closing my discussion of *si*, I should say a few words about usage, for there are certain possibilities and restrictions in this area that require comment. In general, these may be summed up by the following two statements:

1) If a speaker wishes to express deference or formality, he is restricted to using the variant /si/, almost always followed by the deferential particle /khá/ (woman speaking) or /khráp/ (man speaking).

2) There are several kinds of usage of *si* that could be termed assertive, including those that express demandingness, hostility, opposition, rebuke, correction, and the like. Such usage is ideally restricted to speech with intimates or inferiors; and the stronger the assertive element the tighter the restriction.

From the first statement, above, we may draw the inference that all variants of *si* except /si/ suggest a certain amount of informality or familiarity. And this is not surprising, for formal or deferential situations are ones in which we would expect Thai culture to prescribe a certain amount of distance or non-involvement. And informal or familiar situations are ones in which we would expect a speaker to feel



free to express things like definiteness, need or desire, persuasion, and the like - things that are conveyed by the use of the formally developed forms of the particle.

The second statement above presents a problem in that most variants of *si* can be either more or less assertive, depending on the situation. Situations or usages implying little or no assertiveness include invitations, suggestions, simple requests, straightforward responses to questions and question-raising statements, most instances of statements noting new information, and probably all inferential statements. In such situations the form in question may be used rather freely in speaking to almost anyone with whom particular deference is not required. Such usage need presuppose no very close intimacy in use to equals; and it may occasionally occur in speech to intimates slightly superior to the speaker, provided the relationship is a free and easy one.

On the other hand, certain other usages do imply a certain amount of assertiveness. These include the following:

1) All occurrences of /!sî/ and /!sîi/ (i.e. raised /sî/ and /sîi/), and also /s'î/. It is my impression that, of these forms, /!sî/ tends to be more assertive than the others; for the element of persuasion in the other forms softens the element of flat rebuke, contradiction, or hostility that tends to be present in comparable utterances where /!sî/ occurs.

2) Any occurrence of /sîi/ which expresses the hostile type of non-involvement. (See examples 36-39 and subsequent discussion under 3.3.2.1.)

3) Any occurrence of /s'î/ in which the speaker makes socially excessive demands upon the addressee. Now all uses of /s'î/ express some demand that is made of the addressee; for the form by definition involves some wish or need that the speaker calls for the addressee to meet. The crucial question here is whether the demand is excessive or not; and this in turn depends upon the speaker's relative superiority-inferiority and/or intimacy with respect to the addressee, and also upon the nature of the request made. Thus, for example, a superior can use /s'î/ in asking an inferior to run an errand for him - even one involving considerable effort and inconvenience - without necessarily coming across as overly demanding. But in speaking to an intimate equal, a speaker must make rather lesser demands if he does not want to provoke a negative reaction. Thus he can ask the addressee to reach something on a shelf too high for him (the speaker), or to close a window close by the addressee, or to perform some simple service that

the latter can perform more readily than the speaker can. And when the speaker uses /sí/ in such circumstances, his request will come across as a natural thing between intimates - even in certain cases where the addressee may be the superior. On the other hand, if the request involves real inconvenience to the addressee or calls for a service that the speaker could just as easily perform for himself, then the use of /sí/ will probably be taken as overly demanding and assertive, even in speech to intimate equals. And, of course, the likelihood of being so taken will be even stronger if the relationship is not an intimate one.

Given the status-formality restrictions upon the use of variants of *si*, we can readily see that polite words like /chəən/ 'please', or /karunaa/ 'be gracious (enough to)', will not be expected to co-occur with assertively used forms of *si*, and seldom with any of the developed forms of the particle. They may, however, occur with /sí khá/ or /sí khráp/.

## 7. CONCLUSION

I have now carried my treatment of *si* just about as far as I wish to carry it in this paper. In conclusion, however, let me present a summary of the facts that I have set forth above; and then I shall suggest a couple of matters that will eventually demand consideration if we are to achieve a reasonably complete understanding of phenomena relating to *si* and other particles.

### 7.1. SUMMARY OF PHENOMENA PRESENTED ABOVE

The data and conclusions set forth above may be summarised in terms of the following generalisations:

1) There is a particle *si* which, in all its variations of form and meaning, conveys the basic idea of a given response being the logical, necessary, expectable, or appropriate one under the circumstances. This form, with its basic meaning as stated, may occur in action-inducement utterances (commands, suggestions, requests, invitations), in responses to questions and to question-raising statements, and in statements that make an inference or call attention to something.

2) Modifications of length and pitch give rise to the following variant forms with their concomitant semantic values or implications, these values or implications being added then to the basic semantic value of *si* as stated above:

/si/ the speaker is not personally or emotionally involved in the response in question, but he is not particularly trying to call attention to that fact. This is also the neutralised form which occurs whenever si is immediately followed in the utterance by another particle.

/sî/ the speaker is definitely, though unemphatically, involved in the response, there being no ambiguity, doubt, diffidence, in his utterance.

/sîf/ the speaker wants or needs the addressee to do something.

/sii/ the speaker is not personally or emotionally involved in the given response, and he is making a point of conveying this non-involvement, either as a simple matter of fact or as an expression of withdrawal or hostility.

/sii/ the speaker is persuading the addressee to act or to accept the speaker's response.

/sîi/ the speaker urgently wants or needs the addressee to do something and is applying pressure, persuasion.

3) The forms /si/ and /sîi/ can be 'raised' (i.e. raised in pitch so that the falling tone begins above the mid-tone pitch level) to express greater emphasis, definiteness, intensity.

4) The form si is subject to neutralisation when followed immediately by another particle. That is, all potential variants are short-circuited so that only the form /si/ may occur.

The above information can be summarised formulaically as follows:

(1) Differentiation Rule:

$$si<exp> \rightarrow \left[ \begin{array}{l} /si/<exp. + non-in.> (+ length<int.>) \\ /sî/<exp. + def.> \quad (+ length<per.>)(+ raising<emp.>) \\ /sîf/<exp. + s.w.> \quad (+ length<per.>) \end{array} \right]$$

(2) Neutralisation Rule:

$$si \rightarrow /si// - \text{Particle}$$

In the above formulae, pointed brackets <> indicate semantic values (exp. for expectable response, non-in. for non-involvement, int. for intentionality, def. for definite, per. for persuasion, emp. for emphasis, s.w. for speaker's wish); parentheses () indicate optional elements; and square brackets surrounding vertical listings indicate that either one or other of the vertically listed alternatives will occur.

## 7.2. FURTHER AND WIDER CONSIDERATIONS

The above summaries, both as presented in prose and formulaically, imply certain analytical and even theoretical conclusions which I have made no attempt to justify, apart from such justification as may be involved in accommodating my analysis to the facts as I perceive them. Nor do I intend to present such justification here, for to do so would involve a consideration of matters that go far beyond the scope of this paper. But, as I mentioned, I do wish to call attention to a couple of matters related to *si* which involve much more general phenomena in the language. These will eventually have to be considered in depth before anyone can claim to have accounted reasonably adequately for the behaviour of *si*, and before my analysis above can be justified.

Note, for example, the following facts, some of which have already been referred to above:

1) There are certain phonological characteristics which *si* shares with some of the other discourse particles, but not with other forms in the language. These include the lower-than-normal pitch of falling-tone forms and their potential for raising under conditions of emphasis, the absence of terminal glottal stop in short-vowel forms, and the tendency toward a more than normal fuzziness in certain phonemic distinctions.

2) The variants of *si* as described throughout this paper seem to signal semantic distinctions that in some respects appear much more like intonational distinctions than anything else. That is, we can vary pitch and vowel length (within certain limits) and still come up with alternate forms that mean more or less the same thing. Obviously one cannot do this with other forms in the language (such as /*thii*/ '*occasion*', or /*mí*/ '*not*'). But one can do this (again within certain limits) with some of the other particles.

3) In comparing *si* with other discourse particles, particularly in the light of the phenomena just mentioned above, we find that one of them (the particle *na*, speaker's question, wish, or demand) is very like *si* in many ways. A number of others (the status-intimacy particles /*khá*/-/*khâ*/, /*cá*/-/*câ*/, etc.) form a group that are somewhat like *si* in terms of formal variability and other phonetic characteristics, but less so than *na*. And others are hardly like *si* at all.

If phenomena such as the above are to be explained, it would be desirable to do a careful study of all the discourse particles, both individually and as a class - individually to pinpoint and explain variations in form and meaning for each particle, and as a class to see

what parallels may exist between the various particles, and to find out what general statements may be made about the class as a whole or about subgroupings within the class.<sup>9</sup>

Then it would be helpful, further, to examine general phenomena relating to tone, vowel length, stress, and intonation, to see if some new perceptions and correlations can be brought to bear to explain the special characteristics of *si* and other particles.

These two tasks - a general study of particles and a consideration of tone, stress, intonations, etc. - I intend to pursue as I am able. Meantime I here offer my current findings on the forms and meanings of *si*. I hope they will serve both as useful information in their own right and as a starting point for further studies on it and other particles.

## N O T E S

1. Throughout this paper, the representation *si* is used whenever I wish to refer to the particle in general, irrespective of the particular phonological value it may have in a particular instance. Forms cited between slashes are phonemic representations transcribed in accordance with the system of phonemicisation used by Marvin Brown (1967, *et al*). This system is in turn an adaptation of that developed by Mary Haas (1964, *et al*). For a comprehensive and very helpful summary of the various transcription systems currently in use, see Palmer 1974:xvii-xxi.

The values of the transcription symbols used in this paper may be summarised briefly as follows:

Consonants: /p, t, c, k/ are voiceless, unaspirated stops, the /c/ being also affricated; /ph, th, ch, kh/ are their voiceless, aspirated counterparts; /b, d/ are voiced stops; /f, s, h/ are voiceless spirants; /m, n, ŋ/ are voiced nasals; /w, ɣ/ are voiced semivowels; /l/ is a voiced lateral; /r/ is a trilled or flapped, voiced retroflex; and /ʔ/ is a glottal stop.

Vowel combinations: /i, e, ε/ are front, unrounded vowels, high, mid, and low, respectively; /ɤ, ə, a/ are central, unrounded vowels, high, mid, and low; /u, o, ɔ/ are back, rounded vowels, high, mid, and low. All nine vowels may be either short or long - the latter being represented by geminate symbols (/ii/, /ee/, etc.). Diphthong combinations comprise the following: /ia, ɤa, ua/, /iw, ew, eew, εw, εεw, aw, aaw, iaw/, and /uy, ooy, ɔy, ɔɔy, ay, aay, əəy, ɤay, uay/.

Tones are: mid (no symbol), low /`/, falling /˘/, high /˘˘/, rising /˘˘/. On a scale numbered from 1 to 5 (1 being the lowest pitch level, and 5 the highest), the approximate pitch values of the five tones are 33, 22, 42, 44, and 24, respectively.

2. The two final particles in this sentence are sometimes pronounced /na sî/, sometimes /nâ si/, and perhaps occasionally /nâ sî/, with no difference in meaning between the three. The pronunciation /nâ si/ is probably the most common, but for the sake of consistency the /na sî/ variant is given here.

3. It is only fair to note, however, that Brown, in his definition, is only seeking to account for a particular usage of /sî/, not for the whole range of occurrences. His definition does adequately account for the facts he is dealing with.

4. The same is probably true of most, perhaps all so-called sentence-final particles. See for example, the particles /nâ/ (particles marking old information) and /lǎə/ (confirmation particle marking a yes-no question) in the following: /'aacaan<sub>1</sub> nâ lǎə kròot<sub>2</sub>/ 'You mean the PROFESSOR<sub>1</sub> (of all people) got angry<sub>2</sub>?'.  
*PROFESSOR<sub>1</sub> (of all people) got angry<sub>2</sub>?'.*

5. Ms Kiranand informs me that there is no difference in meaning between /si/ and /sii/. However, I am fairly sure that when a speaker really wants to convey unambiguously the meaning implied by the use of either variant, he will choose /sii/, partly because it is easier for the hearer to identify, and partly because /si/ might be taken as a semantically non-differentiated or neutralised form. The phenomenon of neutralisation will be discussed below.

6. See note 5.

7. High tone forms might appear to undergo raising, but actually when /sí/ or /síi/ are raised, the whole utterance must be raised. For example in the utterance /láaŋ sí/ 'Wash it, would you?', the /sí/ cannot be raised significantly higher than /láaŋ/.

8. Noss, however (1964:210), indicates that the statement form /câ/, /khâ/, etc. is possible. I have been unsuccessful in eliciting such an occurrence; but one of my assistants has informed me that some speakers of an older generation might use statement forms of these particles following /si/. I am unable to account for the fact that it is the question form of these particles that usually occurs; for si in none of its occurrences really signals a question.

9. As a matter of fact, Henderson (1949), Chuenkongchoo (1956), and others, have already attempted to make statements about particles in general; but in my opinion these statements have proved comparatively unrevealing because they are based on insufficient data concerning individual particles. For example, Chuenkongchoo makes statements about what certain prosodic combinations (such as short vowel with high tone, or long vowel with falling tone) mean when they occur in particles. He also gives examples of each combinations for each particle. But he never tells the reader what each variant of a given particle means; and, as it turns out, a number of his generalisations do not work in particular cases.



BIBLIOGRAPHY

BHAMORAPUT, Amara

1972 Final Particles in Thai. M.A. thesis, Brown University.

BROWN, J. Marvin

1967-69 *A.U.A. Language Center Thai Course*. Books 1, 2, and 3.  
(Book 1, Bangkok 1967, Book 2, Bangkok 1968, Book 3,  
Bangkok 1969).

CHUENKONGCHOO, Terd

1956 The Prosodic Characteristics of Certain Particles in Spoken  
Thai. M.A. thesis, University of London.

HAAS, Mary R.

1964 *Thai-English Student's Dictionary*. Stanford, California:  
Stanford University Press.

HENDERSON, E.J.A.

1949 'Prosodies in Siamese: a Study in Synthesis'. *Asia Major*  
1/2:189-215.

McFARLAND, George Bradley

1954 *Thai-English Dictionary*. Stanford, California: Stanford  
University Press.

MĀNIT MĀNITCHARĒN

1964 *Photchanānukrom Thai* [Thai Dictionary]. Bangkok: 'Ēksin  
Kānphim.

NOSS, Richard B.

- 1964 *Thai Reference Grammar*. Washington, D.C.: Foreign Service Institute.

PALMER, Adrian S.

- 1974 *Small Talk*. A.U.A. Language Centre Thai Course, Dialog Book A. Bangkok.

RUDARAVANIJA, Panninee

- 1965 An Analysis of the Elements in Thai that Correspond to the Basic Intonation Patterns of English. D.Ed. dissertation, Columbia University.

THAILAND: Rāṭṭhabanditsathān [Royal Academy]

- 1976 *Phoṭṭhanānukrom Chabab Rāṭṭhabanditsathān* [Thai Dictionary, Royal Academy.] Bangkok.

UDOMPHOL, Nisa

- 1969 Semantic Functions of the Thai Particle /si/. Manuscript.

## PHONOLOGICAL REDUCTION OF SOME FINAL PARTICLES IN MODERN THAI

PATCHARIN PEYASANTIWONG

In spoken Thai, particles are used a great deal, especially in casual speech; it seems that the more casual the utterance, the greater the use of particles. Some particles, such as the following, occur before the verb.

kô	as in:	chán kô khít yaŋŋán mǎnkan I pt think like that also 'I think so too.'
ca	as in:	khǎw ca pay mǎn thay phrûŋnǐi he pt go Thailand tomorrow 'He is going to Thailand tomorrow.'
mây	as in:	thǎa mây thoo maa chán kô mây pay if not call come I pt not go 'If you don't call, I won't go.'
cəŋ(thǎŋ)	as in:	thǎe tǔŋ thoo maa chán thǎŋ ca pay you must call come I pt will go 'You must call, then I will go.'

Most particles, however, whether they occur in phrases, clauses or sentences, appear in final position, as in the following examples:

rǎ	as in:	thǎe mây pay thamŋaan rǎ you not go work Q pt 'Aren't you going to work?'
sǐ	as in:	pay sǐ go pt 'Yes, I am.' (Of course I am.)
thǎ?	as in:	yùu bǎan thǎ? stay home pt 'Please stay home.'

læy	as in:	yàa pay læy don't go pt 'Don't go.'
lâ	as in:	thammay lâ why pt 'Why?' (Why not?)
nòy	as in:	yùu pen phuan chán nòy stay be company I pt 'Please stay and keep me company.'
ròk	as in:	mây dâay ròk chán tòn thamnaan not can pt I must work 'I can't, I have to work.'
ná	as in:	tháa yangnán chán pay káp theə ná if like that I go with you pt 'In that case can I go with you?'
sĩa	as in:	tháa ca pay kô pay tèn tua sĩa if will go pt go get dressed pt 'If you want to go, then go get dressed.'
		etc.

Most of these final particles are used in utterances to reveal the speaker's attitude, thought, emotion, mood, etc. regarding the situation at the time of speaking.

The final particles which will be discussed in detail in this paper share a common pattern of phonological reduction of initials, finals, vowels and tones; they can, but do not always, reduce to a common form a, with or without final glottal stop (?), and with one of the three tones, mid, low, or low falling.

The particles to be discussed here are *นะ* /nâ?/, *ละ* /lá?/, *ล่ะ* /lâ?/, *แหละ* /lè?/, and *หรอก* /ròk/. The particle *นะ* /ná?/, while it does not undergo consonant, vowel, or tone reduction, is also included in this discussion because of its close semantic relationship with *นะ* /nâ?/, which does undergo reduction.<sup>1</sup>

Mary Haas (1964) gives a general description of these particles in her *Thai-English Student's Dictionary* which I have taken into account and to which I have added some detail. I have also taken into account Richard Noss's mention of these particles in his *Thai Reference Grammar* (Noss 1964).

We will now look at the way each of these particles is used in utterances, along with examples, most of which are taken from spontaneous conversation. Following this, we will examine the phonological reduction mentioned in the preceding paragraph. Note that reduction

to a with one of the tones mentioned above is not possible in all functions of a particle; in some cases only partial reduction occurs. Where complete reduction is possible, the example is marked with the symbol #.

# PARTICLE USE

๖๖ /nâ?/ is used in sentences:

1. Indicating urging, insisting, persuading, or importuning, such as in the situation in which one is invited to go somewhere and answers rather negatively, or shows reluctance to accept the invitation, the inviter might say:

- (1a) pay /nâ?/  
go pt  
'(Come on!) Please go.'

or in persuading one's guest to stay as in:

- (1b) yuù kòon thà /nâ?/  
stay before pt pt  
'Please stay for a while.' (Don't go yet.)

2. Expressing the speaker's belief,<sup>2</sup> which, the speaker realises, contrasts slightly with the addressee's belief, as in:

- (2a) chán wàa kháw maa lɛ́w /nâ?/  
I say he come already pt  
'I believe he has already come.'
- (2b) thâu-níi ca phoo rɔ́w / / khon phoo /nâ?/  
this much will enough Q / surely enough pt  
'Will this be enough? / It should be.'

3. Marking the topic of that which immediately follows. The topic, which precedes the particle, can be a simple noun, a phrase or a clause (short or long), as in:

- (3a) nən /nâ?/ khaykhay kô yàak-dáay  
money pt anybody pt want  
'Money! Anybody wants it.'
- (3b) phúu-yǐn /nâ?/ tɛn-tua cháa thǐsút  
girl pt get dressed slow most  
'Girls! They get dressed the slowest of all.'
- (3c) tɔon thəə pay dii-sii /nâ?/ chán mây-sabaay  
when you go Washington, D.C. pt I not well  
'When you went to Washington, D.C., I was sick.'

Sometimes the topic comes after the main clause and functions as a declaration. Some people call this an afterthought topic, as in (3d). Examples (3a), (3b) and (3c) could also be rewritten or spoken in this form as well.

- (3d) mây mii kh-ray maa kàp khun ləəy r<sup>h</sup> thii sòp khāw dāy  
 not have person come with you at all Q that test enter able  
 /nâ?/  
 pt  
 'Didn't anybody come with you, from among those who passed the entrance exam?'

4. Indicating the speaker's annoyance. This is used with firm tone of voice.

- (4a) yùt r<sup>h</sup> sǎa-thii thà /nâ?/  
 stop cry pt pt pt  
 'Please stop crying.' (Would you please stop crying?)
- (4b) ruú ləəw /nâ?/, mây tōn maa bōk sámsám-sáaksák  
 know already pt not must come tell repeatedly  
 'I know. You don't have to tell me over and over again.'

5. Indicating response of limited concession, acceptance or agreement on a result. This type of statement usually has the pattern 'verb<sub>1</sub> /nâ?/ verb<sub>1</sub> + clause' or 'verb<sub>1</sub> /nâ?/ neg. verb<sub>1</sub> + clause'. The clause that follows shows that there is some contradiction within the statement, and the pattern is used to show that the speaker is aware of the contradictory nature of his utterance.

- (5a) ruú /nâ?/ ruú, tēə tham ʔaray mây dāy  
 know pt know but do anything not able  
 'I know, but there is nothing I can do (about it).'
- (5b) hēn /nâ?/ mây hēn rōk, tēə ruú wā khāw pen khon ʔaw-pay  
 see pt not see pt but know that he be person take  
 'I know I didn't see him taking it but I'm sure he is the one (who took it).'

6. Ending the utterance with some focus or emphasis, mostly in questions as in:

- (6a) (ca) pay năy /nâ?/  
 will go where pt  
 'And just where are you going?'
- (6b) maa thamay /nâ?/  
 come why pt  
 'And just what are you here for?'

ư /náʔ/ is used in sentences:

7. Making the utterance milder and more gentle, as in declarative sentences, which convey a mild command, or request for an affirmative answer, or at least for acknowledgement from the addressee. For example:

- (7a) pay lá /náʔ/  
go pt pt  
'I'm going now, O.K.?'  
(7b) năngsǔu yuù bon tóʔ /náʔ/  
book be on table pt  
'The book is on the table, O.K.?'  
(8a) chán wâa mây châi /náʔ/  
I say not be pt  
'I don't think it is, do you?'  
(8b) tɛ̀ chán wâa châi /náʔ/  
but I say be pt  
'But I think it is, don't you?'  
(8c) tɛ̀ welaa phûut kàp faràn khâw phûut dii /náʔ/  
but time talk with westerner he talk nice pt  
'... but when he talks with Westerners, he talks pleasantly.'

8. Making milder an utterance which is contradictory or argumentative, used in both negative and positive statements. For example:

- (8a) chán wâa mây châi /náʔ/  
I say not be pt  
'I don't think it is, do you?'  
(8b) tɛ̀ chán wâa châi /náʔ/  
but I say be pt  
'But I think it is, don't you?'  
(8c) tɛ̀ welaa phûut kàp faràn khâw phûut dii /náʔ/  
but time talk with westerner he talk nice pt  
'... but when he talks with Westerners, he talks pleasantly.'

9. Making imperative mood more gentle, both in the negative and positive. In this usage the tone of voice is rather soft.

- (9a) yàa phâŋ khâw maa /náʔ/  
don't just enter come pt  
'Please don't come in yet.'  
(9b) pìt pratuu duây /náʔ/  
close door also pt  
'Close the door too, please.'

10. Indicating mild questions, asking for opinions. This is as if the speaker proposes some idea or says what he/she thinks is true and asks for agreement from the hearer. This may be rendered in English by the question tag, as in:

- (10a) wanníʔ ʔakaat dii /náʔ/  
today weather nice pt  
'The weather is nice (today), isn't it?'

- (10b) phãa nĩi suây /ná?/  
 cloth this pretty pt  
 'This material is pretty, isn't it?'

There is also the pattern of '/ná?' + addressee + /ná?/' usually at the end of an utterance, in seeking agreement. For example:

- (10b') phãa nĩi suây /ná?/ cɔɔn /ná?/  
 'This material is pretty, John, isn't it?'

11. Indicating a request for repetition of information already given, as in:

- (11a) ʔaray /ná?/  
 what pt  
 'What?' (Say it again. I didn't hear you.)
- (11b) raân năy /ná?/  
 shop which pt  
 'Which shop is it?'
- (11c) khăw ca pay mư̄arây /ná?/  
 he will go when pt  
 'When is it he's going?'

12. Expressing mild reproach and/or criticism, mostly in the form of questions.

- (12a) thammay cháa yaàn nĩi /ná?/  
 why slow like this pt  
 'How come you're so slow?'
- (12b) tham yaàn nĩi ʔiik lɛ̄w /ná?/  
 do like this again pt pt  
 'You've done it again, haven't you?'

13. Expressing displeasure, disappointment or complaint. Generally the speaker repeats the subject of the sentence (addressee) after /ná?/. In these sentences /ná?/ occurs following an initial vocative. For example:

- (13a) khun /ná?/ khun, mây pay hāa chán bả̄n lāy  
 you pt you not go see I some pt  
 'You! You never went to see me at all.'
- (13b) khun /ná?/ khun, thammay tham yaàn nĩi  
 you pt you why do like this  
 'You! Why did you do that?'



14. Indicating statement of invitation as in:

- (14a) yen ni'i pay thaen kháaw thii baan /ná?/  
 evening this go eat rice at house pt  
 'Why don't you come to dinner at my house this evening? Please come ....'
- (14b) pay nân rôt lên kan /ná?/  
 go sit car play together pt  
 'Let's go for a ride, O.K.?'

15. Expressing coaxing, suggesting and requesting, or mild insisting as in:

- (15a) khun tông maa /ná?/  
 you must come pt  
 'You must come, please.'
- (15b) yàa phàn pay /ná?/  
 don't just go pt  
 'Please, don't go yet.'

16. Indicating minor annoyance or threat. Such sentences are said in a firm tone of voice. The imperative quality of the utterance is stressed.

- (16a) yàa kháw maa /ná?/  
 don't enter come pt  
 'Don't come in.'
- (16b) ọĩap /ná?/  
 quiet pt  
 'Be quiet!'

In a different situation the same construction could be a request for an opinion (as in example 10). For example, if spoken in a very quiet place number (16b) would be translated as 'It's quiet, isn't it?'. In such a situation the tone of voice would be a mild one.

17. Marking topics. As with ư /nâ?/ in example (3) above, the topic can be a simple noun, a phrase or a clause. However, with ư /ná?/ the item marked as topic is also given some emphasis. For example:

- (17a) rưn ni'i /ná?/, chán wãa mây còp ròk  
 matter this pt I say not end pt  
 'I don't think I'll ever see the end of this problem.'
- (17b) chán khít /ná?/, wãa mii klòn plaàw yuù baan  
 I think pt that have box empty loc.pt some  
 'I think I have some empty boxes.' (I think there are some empty boxes.)

ละ /láʔ/ and ละ /lâʔ/

The two particles /láʔ/ and /lâʔ/ are used extensively in modern Thai, and there is some evidence to suggest that they are derivations of the particles แลว /leʔw/ and เลว /lêw/. With regard to semantics, the particles /láʔ/ and /leʔw/ are virtually interchangeable, as are the particles /lâʔ/ and /lêw/. The words /leʔw/ and /lêw/ are very common in older works of prose fiction, but /láʔ/ and /lâʔ/ are to be found only very infrequently. Modern prose, on the other hand, makes frequent use of these latter two words. In addition, modern casual conversation relies almost exclusively on /láʔ/ and /lâʔ/. It is also worth noting tonal correspondences between the particles /leʔw/ and /láʔ/ and /lêw/ and /lâʔ/. Finally, George McFarland's (1954) *Thai-English Dictionary*, compiled prior to 1937 includes only /leʔw/ and /lêw/, making no mention of either /láʔ/ or /lâʔ/, whereas Haas (1964) includes all four forms.

Thus there is evidence to suggest that /láʔ/ and /lâʔ/ have only recently become the preferred forms and have been derived from /leʔw/ and /lêw/ by a process of reduction. Although the evidence is not absolutely conclusive, I feel that it is strong enough for the purpose of this paper.

ละ /láʔ/ (as well as แลว /leʔw/) is used in sentences:

18. Indicating that the state of being which is posited by the verb-adjective has just been attained as in:

(18a) phəw /láʔ/  
enough pt  
'That's enough.' (O.K. Enough!)

(18b) dii /láʔ/  
good pt  
'That's good enough.'

The expression ดีละ /dii láʔ/ is sometimes used in a sarcastic sense as in:

(18b') dii /láʔ/, fàak wáy kòn  
good pt entrust keep first  
'Fine, just wait till it's my turn.'

19. Indicating the speaker's consent to something done or said by others as in:

(19a) thùuk /láʔ/ or chây /láʔ/  
correct pt be pt  
'That's correct.' 'That's right.'

20. Situation closing, in order to tell the addressee something and to indicate some change in the situation. For example:

(20a) chán pay /lâʔ/  
I go pt  
'I'm going now.' (I'm going to leave now.)

(20b) chán ca kin /lâʔ/  
I will eat pt  
'I'm going to eat now.'

(20c) chán chák-ca chóp khên maa /lâʔ/  
I start to like ascend come pt  
'I'm starting to like it now.'

21. Indicating a little annoyance, again with certain tone of voice as in:

(21a) maa ʔiik /lâʔ/  
come again pt  
'Here he comes again.'

(21b) ca tham ʔaray kô tham pay thêʔ chán mây sôncay /lâʔ/  
will do what pt do go pt I not interest pt  
'Do whatever you want; I don't care any more.'

ลဲ /lâʔ/ (as well as เล้า /lâw/, mostly used in question form) is used in sentences:

22. Indicating surprise or the unexpectedness of a situation through the use of questions as in:

#(22a) sawàt pay năy /lâʔ/  
Sawat go where pt  
'Where has Sawat gone?'

#(22b) thammay kháw mây maa /lâʔ/  
why he not come pt  
'How come he didn't show up?'

23. Indicating the speaker's irritation or annoyance, also used with irritated tone of voice as in:

#(23a) maa thammay ʔiik /lâʔ/  
come why again pt  
'What are you here for this time?'

#(23b) ʔaray ʔiik /lâʔ/  
what again pt  
'Now what!' (What is it this time?)

24. Referring the question back to the other. This can be rendered in English as 'And you?' or 'What about you?'. For example:

- #(24a) chán kin lɛ́ɛw , khun /lâ?/  
 I eat already you pt  
 'I've already eaten, and you?' (... , what about you?)

Sometimes it is used as a retort as in:

- #(24b) thəə ʔeɛŋ /lâ?/, sǔ̀sàt khě̀-ny  
 you self pt honest how much  
 'And what about you yourself, how honest are you?'

25. Indicating an additional question within conditional statements as in:

- #(25a) thāa khǎw maa /lâ?/  
 if he come pt  
 'What if he should come?'
- #(25b) thāa mây thuuk /lâ?/  
 if not correct pt  
 'What if it's not correct?'

26. Indicating some sarcasm (with sarcastic tone of voice), although not very strong as in:

- (26a) ʔɔ̀! dii nǐi /lâ?/  
 excl. nice pt pt  
 'Oh, nice!' (Oh, swell!)
- (26b) ca bòn ʔiik máy /lâ?/  
 will complain again Q pt  
 'Well, are you going to complain any more?'

Another particle that seems to become laʔ (with mid, rather low or low tone) is the particle ɛ̀lɛʔ. The sound of this particle /laʔ/ just mentioned does not have a common written form but is sometimes written as ɛ̀lɛ or ɛ̀lɛ.<sup>3</sup> ɛ̀lɛ /lɛʔ/ is used in sentences:

27. Intensifying the statement or pointing out something certain and specific. It is usually preceded by the demonstrative nǐi, nǐi (= 'this') or nân, nân (= 'that') as in:

- (27a) rew-khâw, lèn yuù nân /lɛʔ/  
 hurry up play loc.pt dem. pt  
 'Hurry up! Stop fooling around.'
- (27b) chán ʔeɛŋ /lɛʔ/  
 I self pt  
 'It's just me.'

The last particle that will be examined here is *wan* /*rɔ̀k*/. This particle can always be pronounced as *rɔk*, *lɔk*,<sup>4</sup> or *ɔk* (with either mid or low tone) which phonologically seem to be first and second step derivation (1. vowel shortening, 2. *r* → *l*, 3. initial dropping). Very often, especially in fast speech, however, it derives to *àʔ*.

*wan* /*rɔ̀k*/ is used in sentences:

28. Indicating mild negative statement. (This is the most common use of this particle.) For example:

(28a) *chán mây pay /rɔ̀k/*  
*I not go pt*  
*'I'm not going.'*

(28b) *mây sanùk /rɔ̀k/, yàa pay ləəy*  
*not fun pt don't go pt*  
*'It's not fun, don't bother going.'*

29. Expressing reassurance as in:

(29a) *thəə khon ca sɔ̀p pha'an /rɔ̀k/*  
*you surely will take test pass pt*  
*'I'm sure you'll pass the test.'*

30. Expressing mild sarcasm (with sarcastic tone of voice) as in:

(30a) *khon sanùk /rɔ̀k/*  
*surely fun pt*  
*'That'll be fun!'*

(30b) *khon dāy khən ɲəndəan /rɔ̀k/*  
*surely get increase salary pt*  
*'You'll probably get promoted.'* (Keep on sitting here playing cards.)

If the sentences in (30) are not said with a sarcastic tone of voice, they would be like (29) or be used as consolation. In (30a) then may imply something like *'Don't worry about it, just go. It may be fun.'*

31. Indicating some kind of hesitancy in agreeing with somebody as in:

(31a) *kɔ̌ suəy dii /rɔ̀k/*  
*pt pretty good pt*  
*'Yes, it's pretty.'*

It can also be added at the end of the first group in number (5), and make the acceptance a little milder as in:

(31b) *ruú /nàʔ/ ruú /rɔ̀k/ tɛ̀ tham ʔaray mây dāy*  
 (5a) *'I know, I know, but there is nothing I can do about it.'*

32. Indicating a little annoyance or reproach in a sarcastic manner as in:

- (32a) chǎn ʔàan ʔeen pen /rɔ̀k/  
 I read self able pt  
 'I can read it myself (you don't have to read it out loud).'
- (32b) chǎn mii taa duu ʔeen /rɔ̀k/  
 I have eye look self pt  
 'I have my own eyes, you know.'

Note that all the particles /nâʔ/, /náʔ/, /láʔ/, /lâʔ/, /lèʔ/, and /rɔ̀k/ in the examples are presented in their fullest forms. In addition, it should be noted that in each example the particle is placed in utterance-final position. The following part of the study examines the possible phonological reductions that each particle can undergo. Variants of these reduced forms which appear when they are followed by other final particles will be examined as well.

#### DERIVATION OF THESE FINAL PARTICLES

The phonological rules that seem to be involved in the derivation from /lɛ́w/ to /láʔ/ are:

- a) vowel shortening: lɛ́w → lɛ́w  
 b) vowel shifting:<sup>5</sup> lɛ́w → láw  
 c) glottal stop replacing final: láw → láʔ

From /lâw/ to /lâʔ/, only rule (c) applies, since the vowel of the original form is short and low (no rule (a): VV → V, nor rule (b):  $\begin{smallmatrix} \check{V} & + & V \\ [-\text{central}] & [+central] \end{smallmatrix}$ ). As for /lèʔ/ being reduced to /làʔ/, only rule (b) is necessary since the vowel was originally short and the final was the glottal stop.

The following are the reductions that occur in the initial vowels, tones and finals of the particles discussed above.

/nâʔ/	is reduced to	nâ	in 1, 2, 4
"		âʔ	in 3, 5, 6
/náʔ/	is reduced to	ná	in 7-17
(lɛ́w →)/láʔ/	is reduced to	$\begin{Bmatrix} \acute{a}ʔ \\ aʔ \end{Bmatrix}$	in 18-21
(lâw →)/lâʔ/	is reduced to	â	in 22-25
"		lâ	in 26

(lɛ̌ʔ →)/lâʔ/ is reduced to  $\begin{Bmatrix} \text{à}ʔ \\ \text{a}ʔ \end{Bmatrix}$  in 27

/rɔ̌ʔk/ is reduced to  $\begin{Bmatrix} \text{à}ʔ \\ \text{a}ʔ \end{Bmatrix}$  in 28, 29  
 "                      ɔ̌k in 30-32

The chart shows the most reduced form possible for these particles when they are used in sentence final position. The word /náʔ/ always retains the initial but drops the final whereas /nâʔ/ can either retain the initial and drop the final or drop the initial and retain the final, depending upon the expression.

The words /láʔ/ (from /lɛ̌w/) and /lâʔ/ (from /lɛ̌ʔ/) always drop the initial and keep the glottal stop. Their tones can also be leveled towards mid tone. However, /lâʔ/ (from /lâw/) tends to drop both the initial and the final leaving only the vocalic and the tone /â/, except as in item number (31) where the initial /l/ remains.

The particle /rɔ̌ʔk/ is rather unusual in changing to àʔ or aʔ since the original and reduced forms do not share consonants or vowels, and unlike /lɛ̌w/ (which becomes first /láʔ/ and then áʔ or aʔ), it changes from /rɔ̌ʔk/ directly into 'àʔ'. The process for this reduction involves:

- a) vowel shortening: rɔ̌ʔk → rɔ̌k
- b) vowel shifting:                      rɔ̌k → ràk
- c) glottal stop replacing final:      ràk → ràʔ
- d) initial consonant deleting:              ràʔ → àʔ
- e) tone neutralising:                              àʔ → aʔ

Steps (a), (b) and (c) are the same as those which occur in the reduction of /lɛ̌w/ except that there is no particle /ràʔ/ derived from /rɔ̌ʔk/ as /láʔ/ is derived from /lɛ̌w/.

However, in certain situations the final reduction of this particle is ɔ̌k or ɔ̌k which means that only rule (a) and (d) apply (rɔ̌ʔk → rɔ̌k → ɔ̌k). It is also possible to apply only rule (a) (→ rɔ̌k).

The reduced forms above show great similarity in that they share the vocalic nucleus (a low, central, unrounded vowel). The dropping of the initial in the most reduced form causes difficulty in an attempt to determine the underlying form since Thai allows linking between words. For example, níi and nán ('this' and 'that') are often linked with these particles becoming:

	A	B	C	D
I.	nií + nâ?		→ ní + â?	→ <u>nía?</u>
II.	nií + lâw	→ nií + lâ?	→ ní + â	→ <u>nía</u>
III.	nií + lè?	→ nií + là?	→ ní + à?	→ <u>nía?</u>
IV.	nií + rɔ̀k		→ ní + á?	→ <u>nía?</u>

etc. and

V.	nán + nâ?		→ nán + â?	→ <u>nánnâ?</u>
VI.	nán + lâw	→ nán + lâ?	→ nán + â	→ <u>nánnâ</u>
VII.	nán + lè?	→ nán + là?	→ nán + à?	→ <u>nánnâ?</u>
VIII.	nán + rɔ̀k		→ nán + á?	→ <u>nánnâ?</u>

(/lɛ́w/ and /ná?/ are often used with clearer speech like that in column C.)

In the chart above, column D is the final reduction as uttered in fast speech; column C has the same elements as uttered in speech slower than that of D, and thus the forms remain two syllables. Column B shows another possible derivation (/lâw/ → /lâ?/ and /lè?/ → /là?/) which can also be used in slow and clear speech. The citation form is that in column A whereas D is the most common in speech.

In examples I through IV an open syllable with a long vowel is combined with the final reduction of the particles; in such cases the word preceding the particle undergoes vowel shortening as shown in column C and D. The tone of the final particle seems to be of less weight than that of the preceding word, and is usually neutralised. The tone of the new derivation in column D thus takes the tone of that preceding word.

(Thus, nií +  $\left\{ \begin{matrix} \hat{a} \\ \grave{a} \\ a \end{matrix} \right\} \rightarrow \text{nía?}$ ).

(Other examples are níi + â? → níā?, níi + à? → níā?, etc.)

In the first group, I, II and IV have the same reduced form after the linking process; however, they are used in different situations, for example:

I.	mây chây rɔ̀k, ʔan <u>nía?</u> + mây chây rɔ̀k ʔan nií /nâ?/
II.	ʔan <u>nía?</u> + ʔan nií /lè?/
IV.	mây chây ʔan <u>nía?</u> + mây chây ʔan nií /rɔ̀k/

(Note: mây chây = 'not', negative; ʔan = classifier for things; nií = 'this')



Example I is the expression of an afterthought topic (as in (3d)).

Example II is an answer, confirming that a particular topic is the one being discussed (as in (27)).

Example IV is simply a negative statement with negative particle /rɔ̀k/ at the end which is the most common use (as in (28)).

The second group operates the same way except that the word preceding the particle is a closed syllable. Since the form of the final reductions of the particles has no initial consonant, when the two syllables are put together the final consonant of the preceding syllable becomes an ambisyllabic consonant, that is, it functions both as the final consonant of the preceding syllable and the initial consonant of the following syllable.

Confusion also may arise when one of these particles appears in sequence, either with another member of the same group or with other final particles, such as those indicating degree of politeness or formality, including those cited on the first page here. The glottal stop at the end of the reduced forms above, according to the phonological rule, will be dropped when those forms are followed by other particles. In addition, the tone will sometimes be levelled (towards mid tone), which makes most of the forms sound similar, including those of /náʔ/ and /láʔ/ which usually are more distinctive than the others discussed in this paper. Examples of this are:

<i>I</i>	<i>will</i>	<i>sit</i>	<i>here</i>	+ pt	+ polite pt
chán	ca	nân	thîi	<u>nîi</u> /lɛʔ/	/khâʔ/
chán	ca	nân	thîi	<u>nîaʔ</u>	(without polite particle)
chán	ca	nân	thîi	<u>nîa</u>	<u>khâ</u> (polite particle)

<i>go</i>	+ pt	+ question pt
pay	/láʔ/	/rɔ̀k/
pay	<u>áʔ</u>	(without question particle)
pay	<u>á</u>	lɛʔ (with Q pt which derives to lɛʔ)

Thus the final reduction with the glottal stop at the end only occurs when the particle is in sentence-final position. If another final particle follows it, the glottal stop is always dropped. The most reduced form of these particles, then, seems to be a with different tones and very often mid tone.

N O T E S

1. / / indicates the underlying form or the tentative form according to the spelling. \_\_\_\_ (underline) indicates the actual spoken form of the particle.
2. For a stronger expression of this type sometimes the vowel is lengthened. This also applies in the use of number 1 as well.
3. Rules of Thai orthography make it impossible to represent a mid tone on a short vowel. These two forms น้ะ and ะ would be read in isolation with a low and a high tone respectively. Nevertheless, these forms would be read with mid tone in the proper context.
4. In Bangkok Thai /r/ is often replaced by /l/ in spoken language. Thus, lək is used as well.
5. Vowel shifting: low front vowel → low central vowel (e → ə).  
                                [ə-central]

BIBLIOGRAPHY

HAAS, Mary R.

- 1964     *Thai-English Student's Dictionary*. Stanford, California:  
Stanford University Press.

McFARLAND, George Bradley

- 1954     *Thai-English Dictionary*. Stanford, California: Stanford  
University Press.

NOSS, Richard B.

- 1964     *Thai Reference Grammar*. Washington, D.C.: Foreign Service  
Institute.



## HOW MANY TONES FOR SOUTHERN THAI?

ANTHONY DILLER

Sometimes linguistic questions press themselves on us with all the urgency of common sense, and interpretive detachment is needed to wonder some about the questions themselves.<sup>1</sup> Given a tone language, it might seem reasonable to demand how many tones there are, and to be content only with an unambiguous integer as answer: Mandarin, 4; Vietnamese, 6; Black Miao of Yong-Cong, 8. In this paper we examine a case in which the question is unrevealing unless further qualified.

For the Tai language family (including what may be a few 'cousins') tone-count reports show a range:<sup>2</sup>

Dai, Be (Hainan 'cousins')	- 4
Central Thai; Lao varieties	- 5
Black and White Tai	- 6
Chuang of Ta-sing	- 7
Chuang of Yan-sho	- 8
Kam (a Kuei-chow 'cousin')	- 9

Especially to the north, Tai shows a predilection for the 6-tone system as evidenced by over 40 Chuang varieties, over 40 Pu-yi varieties, Nung, Saek, Sui and others. 5-tone systems are also frequent, and all Tai varieties for which reliable tone reports are available fall somewhere in the 3-to-9 range.

Although tones vary phonetically from place to place, tonal correspondence for much vocabulary is regular (less so for 'cousins'). If we lay aside for the moment several restricted cases, including syllables terminating in final stops, a fairly clear picture can be drawn. Vocabulary in a typical 6-tone language such as White Tai can be sorted into 15 correspondence categories. These are designated a-o in Figure 1 and depend both on tone and on status of initial consonant.

FIGURE 1  
Tai 6-tone system

tone	[+aspirated] [-continuant]	[-aspirated] [-continuant]	[-aspirated] [+continuant]	other designations
1	a	b	c	M0-H0; A1; A1-3
2	d	e	f	M1-H1; B1; B1-3
3	g	h	i	M2-H2; C1; C1-3
4	-	j	k	L0; A2; A4
5	-	l	m	L1; B2; B4
6	-	n	o	L2; C2; C4

Examples: a: phaa<sub>1</sub> 'cliff'; b: kaw<sub>1</sub> 'to scratch';  
 c: vaan<sub>1</sub> 'sweet'; d: phaa<sub>2</sub> 'to split'; e: kaw<sub>2</sub> 'old';  
 f: vaan<sub>2</sub> 'to scatter, to sow'; g: thaa<sub>3</sub> 'to wait';  
 h: kaw<sub>3</sub> 'nine'; i: law<sub>3</sub> 'whiskey'; j: taa<sub>4</sub> 'to paint';  
 k: min<sub>4</sub> 'louse'; l: taa<sub>5</sub> 'landing'; m: law<sub>5</sub> 'to tell';  
 n: taa<sub>6</sub> 'to defy'; o: law<sub>6</sub> 'enclosure'

In varieties of this type correspondence categories a, b and c coincide as to tone. This can be called a case of *tonal coincidence* and represented as /abc/. Figure 1 indicates that in languages like White Tai inherited Tai vocabulary items with aspirated-stop initial consonants are restricted as to tone. Some northern varieties have no aspirates at all. But there is no reason yet to be suspicious about asking or answering the number of tones, and these 6-tone systems can be abbreviated /abc/def/ghi/jk/lm/no/, with actual pitch levels and contours for any particular tone varying from place to place.

Nor is there much problem counting tones as one moves southward to varieties like Lao of Luang Prabang or Central Thai where j, l and n items now take on aspirated initials (as though sifted into the left column of Figure 1). The patterns of tonal coincidence for b and h are altered, and in addition we find cases of *tonal merger* in which homophony can result. This can be indicated by brackets; e.g. if /[im]/ appears then items like the following become homonyms:

law (correspondence category i) 'whiskey'  
 law (correspondence category m) 'to tell'

The net effect of these redistributions (from the point of view of Figure 1) may be a 5-tone system such as that of Luang Prabang Lao, /ac/bjk/e[d1][fm]/gi/hno/, or that of Central Thai, usually cited in the order /bjk/def/h[gl][im]/no/ac/.

Tai varieties below the Kra Isthmus can be called Southern Thai and in general have aspirated initials for *j*, *l* and *m* items as in the 5-tone systems above. However patterns of tonal coincidence and tonal merger are different. For major Southern varieties on the Pacific side of the Malay Peninsula tonal detail shifts areally as represented in Figure 2. In this case latitude is a convenient index and triplets of numbers from 1 to 5 indicate low to high features of pitch level and contour.<sup>3</sup> Note that in all the varieties for correspondence categories *a*, *b*, *c*, *d*, *e* and *f* there are homophonic mergers [*ad*], [*cf*] and [*be*]; also [*ad*] everywhere coincides with [*cf*].

FIGURE 2  
Southern Thai tonal detail

	<i>ad</i>	<i>cf</i>	<i>be</i>	<i>gi</i>	<i>h</i>	<i>jk</i>	<i>lm</i>	<i>no</i>
10°N	532	532	231	333	333	231	234	221
9°N	543	543	343	"	"	331	"	"
8°N	454	454	344	443	"	321	223	211
7°N	455	455	"	"	"	"	222	"

Examples: *ad*: phaa 'cliff', 'to split' (i.e. homonyms);  
*cf*: waan 'sweet', 'to scatter'; *be*: kaw 'to scratch',  
 'old'; *gi*: thaa 'to wait'; laaw 'whiskey'; *h*: kaaw  
 'nine'; *jk*: thaa 'to paint'; mien 'louse'; *lm*: thaa  
 'landing'; laaw 'to tell'; *no*: thaa 'to defy'; laaw  
 'enclosure'

The majority of Southern Thais speak varieties in this general range. Above and in pockets below the zone there are separate characteristics, and on Samui Island *h* coincides with *lm*. The West Coast has a smaller, historically less stable population; Figure 2 is still applicable, but with a measure of areal mixing. Cf. Court (1975) for one such dialect; his Note 5 raises in a general way some of the questions discussed below.

How many tones are there in Southern Thai?

To answer this, clearly we must first distinguish areal subtypes. For a First Approximation discrete contours and pitches in Figure 2 can simply be counted for each areal variety. Near 10°N /*[be]jk/* and /*gih/* are each phonetically single tonal entities, so the 5-tone system /*[ad][cf][be]jk/gih/lm/no/* is distinguished. Farther south there is first a 6-tone, then a 7-tone system. The former has been reported by Jones (1965),<sup>4</sup> and the latter by Haas (1958) for Nakhon Srithammarat and by Henderson (1959) for Songkhla, etc.

If we are counting phonetic tonal entities we need to expose a tacit assumption made above. So far only isolated or phrase-final stressed syllables have been subject to counting. In real speech contexts tone languages are notorious for their tonal sandhi and Tai languages are no exception.

A Second Approximation might be undertaken along this line inquiry. Gandour (1975) has called attention to loss of rising contour in certain Phuket forms (viz. 1m) and Vichin Panupong (1972) has presented measurements for Songkhla pretonic syllables from categories *a*, *d*, *c*, and *f*, which undergo contour reversal (from about 455 to 54 or 43) as do those from *b* and *e* (344 to 32). For Songkhla (7°N) speech Vichin thus enumerated 9 discrete tonal entities, viz. 7 tonic plus 2 pretonic alternates.

Probably 'at least 9' would be a safer count. In Songkhla speech citation-form tonic homonyms occur in the *g* and *i* categories, e.g.

maay-443    'to burn'; 'preverbal negative form'; 'phrase-final question particle'

In slow unnatural articulation the following sequence is ambiguous:

maay-443    maay-443

Interpretations are (a) '*Is it burning?*' and (b) '*It is not burning*'. However in natural rapid speech placement of phrase stress disambiguates (b) by assigning a shortened noncontour alternate for the preverbal negative form.

Court (1975:70) has distinguished a grand total of nineteen etic tones for Satun; this includes canonical forms, free variants and sandhi forms.

Other examples could be cited. The problem is when in principle to stop counting. The more opulent a tone system becomes on the basis of tone sandhi, the more confident one is that for most purposes tone counting should involve notions of contrast and complementarity. One assumes that sandhi can be 'taken care of' by automatic rules of some sort. Lexically contrastive use of tone on the other hand is a basic phonological constituent for tone languages.

A Third Approximation is required to represent the tonal economy of the lexicon. Although some may consider it antediluvian, the tonal phoneme (toneme, distinctive prosodeme, etc.) is here the issue, and data in Figure 2 can be subjected to rather classical structural analysis. We note that items of categories *b*, *e* and *h* have unaspirated stops as initial consonants, and such initials are in complementary distribution with those in other categories. This suggests reduction to 5-tone systems for the 7°N, 8°N and 9°N varieties. The 10°N variety is phonetically



(i.e. by the First Approximation) 5-tone already and cannot be further reduced by this means. Cf. Court (1975) for a discussion of the Satun system and its reduction; Songkhla 7°N is so reduced below.

The actual construction of these reduced systems must rely on a further somewhat unbridled notion, that of phonetic similarity, but now applied to tones. Wang (1967) has made important progress in this direction, and it would appear that close parallelism (perhaps with differentiation by a single nonkinetic feature) is a natural criterion for establishing tonal allophones.

Although the reduced tone systems may be non-unique solutions, they need not be wholly arbitrary. Take the case of 7°N. The problem here is to find for [be] and for *h* corresponding aspirate-initial tones to complete tonal phonemic units. For [be] a good choice is clear: not only does [ad][cf] have a contour both parallel and immediately adjacent, but phonological behaviour is shared as well. In pretonic tonal sandhi we noted above that both [ad][cf] and [be] are subject to the same sort of contour reversal. This is good motivation for formulating a lexical tone, say, tone 1, as /[ad][cf] + [be]/, for which [be], the allophone on lower pitch, occurs in conjunction with unaspirated-stop initials. The foregoing can be summarised in two rules if one follows the notation of Wang (1967) and further allows the *gaucherie* αβ to represent multiplication of variable algebraic signs.

Rule 1

$$\text{tone 1} \rightarrow \begin{array}{l} [\alpha\beta\text{high}] \\ [-\text{low}] \\ [+ \text{contour}] \end{array} \Bigg/ \begin{array}{l} [\alpha\text{aspirated}] \\ [-\beta\text{continuant}] \end{array} \text{ —}$$

Rule 2

$$[+ \text{contour}] \rightarrow [\gamma\text{falling}] \Bigg/ \underline{[-\gamma\text{stress}]}$$

Similarly *gi* can be plausibly paired with *h*. The fact that the resulting system /[ad][cf] + [be]/*gi* + *h/jk/lm/no*/ is equivalent to the 6-tone system of Figure 1 with its tones 1 and 2 merged is interesting diachronically but need not be adduced as evidence for present purposes.

However a distributional addendum does second the solution above. In the 7°-8°N area unaspirated-stop initials sometimes occur with the 'wrong' tones in Malay loans and in other items not regular as to Tai correspondence category. Thus the initials *b-* and *d-* may occur in items coinciding tonally with *jk* or *lm*:

- baŋ-321 'address term to older male' (cp. Malay abang)  
 bæ-222 'since, last'  
 daŋ-321 'wide-weave winnowing basket' (cp. Malay badang)  
 dee-321 'this (emphatic)'  
 doo-321 'that (emphatic)'

Of course this is no challenge to the 5-tone system proposed above since the additional data simply indicate that unaspirated-stop initials, like others, are actually unrestricted as to lexical tone. But other solutions would be strained. If, say, [be] were taken as an alternate of *jk* (perhaps on the basis of the 10°N situation) allophonic conditioning would become problematic and conflicts would occur with forms like:

- baŋ-344 'in the way (visually)'  
 daŋ-344 'loud'

Similarly [be] would be uncomfortable with *lm*.

As one moves north of 7°N cohesive parallelism within /*[ad][cf]+[be]*/ begins to disintegrate and also the addendum above becomes insignificant (through shifted tones, occurrence of nasal initials instead of *b-* and *d-*, and infrequency of Malay loans). [be] eases its way phonetically toward *jk* and finally coincides with it at 10°N as can be seen in Figure 2.

So by the Third Approximation we can enumerate 5 structural tones at 10°N automatically and again at 7°N quite plausibly. Between these systems there appears to be a certain transitional ambivalence: while 5-tone structures can still be imposed on 6- and 7-tone phonetic backgrounds through invoking complementarity, the resulting solutions seem less compelling than the others.

In fact for certain transitional varieties in the Trang-Phatthalung area complementarity also disintegrates. Sporadically for some, systematically for others, aspirated-stop initials are replaced by unaspirated counterparts with segmental functional load shifted to the tonal system.<sup>5</sup>

For clarity, in the preceding approximation certain facts of initial palatal nasal and semivowel distribution have been overlooked. We can now turn to these matters and see how they affect the structural tone counts. For this Fourth Approximation an added dimension of sociolinguistic variation needs to be considered.

'Urban hybrid' as distinct from rural Southern Thai varieties need to be distinguished, although these labels are only suggestive. Structurally an urban hybrid is a mixture of Central Thai segmentals and Southern Thai tones, i.e. the tones of one of the Southern areal systems. Some illustrations are given in Figure 3. Functionally the hybrids constitute the normal colloquial vehicle among educated urbanised

Southerners, although many may control both hybrid and rural (i.e. more 'purely Southern') forms as an important resource for manipulating social distance.

FIGURE 3

Southern Thai (7°N) rural, urban hybrid and Central Thai forms

Southern rural	Southern urban hybrid	Central	
ñiŋ-455	yŋ-455	yŋ-234	'woman'
ñay-455	yay-455	yay-221	'big'
ñaa-443	yaa-443	yaa-432	'grass'
hɔɔn-455	ŋɔɔn-455	ŋɔɔn-234	'cockscorn'
see-455	sii-455	sii-221	'four'
mlæən-222	lɪɪn-222	lɪɪn-432	'slippery'

The occurrence of initial ñ- is restricted areally, socially, and also with respect to tonal distribution.<sup>6</sup> Elsewhere y- occurs in vocabulary items in the [+continuant] categories. It also occurs as though [-continuant] in a rather aberrant and restricted set of items we can designate y\*.<sup>7</sup> Details of distribution are summarised in Figure 4.

FIGURE 4

Distribution of initial palatal continuants in Southern Thai

correspondence categories	7°-8°N rural	7°-8°N urban hybrid	9°-10°N
c, f, i	ñ	y	y
k	ñ, y	y	y
m, o	y	y	y
y*	y (tone be)	y (tone be)	y (tone cf)

How are these segmental particulars relevant to the counting of tones?

If one is inclined to be strict in examining systematic complementarity the curious distribution of y\* especially as it relates to ñ- becomes structurally significant.

At 7°N, for rural speech the 5-tone system of the Third Approximation is confirmed. Rule 1 simply needs to be expanded to specify the higher-pitch allophone of tone 1 for ñ-, the lower-pitch one for y-. (Furthermore an inverse analytic possibility of recognising 7 tones in the lexicon and taking segmental pairs ñ-/y- along with ph-/p-, etc. as allophonic alternates is now weakened; note pairs with characteristic k such as:

ña-a-321 'preverbal negative imperative marker'

yaa-321 '*species of fish (Rastrelliger)*'

However for the 7°N urban hybrid the preceding pair are homonyms and [cf] and [be] no longer have initial consonants in strict complimentary distribution. This means 6 structural tones.

The same can be said for 8°N and holds until a mountainous isogloss signals the phonetic 'leap' of  $y^*$  out of [be] into [cf]. At this point we are back to the 5 structural tones of the Third Approximation. (This 'leap' is diachronically interesting and is the only major Southern areal shift which is not a gradient one. Near the isogloss running from Khanom through Na San south to Krabi-Trang competing  $y^*$  [be]~[cf] alternates are heard.)

The outcome of the above ñ-/y- Fourth Approximation is thus to confirm 5-tone structural systems at ends of a continuum to which a social axis is now added. At the same time the questionable 5-tone systems for intermediate varieties are now technically invalid.

A final Fifth Approximation is needed in order to revise slightly the earlier counts with respect to syllables terminating in final stops. In most Tai cases stopped syllables coincide tonally with one or another of the unstopped categories a-o with no tonal proliferation. Figure 5 shows that this is generally true for Southern Thai, but with a couple of special considerations.

FIGURE 5

Tonalities of representative Southern Thai stop-final syllables

	7°N rural	9°-10°N	cp.	
thɔɔt	444	444	g	'to take off'
laat	444	444	i	'market'
{ ñaap	444	-	i	'rough'
{ yaap	-	444		
yiap	333	444	$y^*$	'to trample'
tɔɔt	333	234	h, l	'to nibble'
thɔɔt	222	234	l	'to cast'
laat	222	234	m	'to spread'

Firstly, in the more northern varieties an extra phonetic (i.e. First Approximation) entity must be recognised, 444, which fails to coincide exactly with one of the unstopped categories. However structurally (i.e. by the Third and Fourth Approximations), the concurrent criteria of complementarity and contour agreement for projected allophones again preserve a system of 5 tonal contrasts.

Secondly, in the varieties farther south no phonetic additions are required, but the palatal continuants  $y$ - and  $\tilde{n}$ - now occur with problematic tonalities 333 (e.g. *yiap* 'to trample') and 444 (e.g. *ñaap*, *yaap* 'rough'). Lexical tone /*gi* + *h*/ is now technically valid only for those varieties consistently distinguishing  $\tilde{n}$ - for *i*-items at 444 (viz. 7°-8°N rural). Elsewhere the previous structural count of 6 tones must be upped to 7, bringing us full circle back to the count of the First Approximation. The count 7 now has structural validity as well, although this rests on the distribution of a single initial consonant.<sup>8</sup>

In conclusion it might be charged that the discussion above is a matter of outdated quibbling, and that contrast and complementarity should be left behind when suprasegmental/segmental relationships are at issue. But not so. Clearly, a valuable result confirms the analytic usefulness even of narrowly-defined complementarity. The result is that independent structural criteria are found to *converge* in certain cases.

At two separated points in an areal-social continuum distinct 5-tone systems become quite cohesive phonologically. Segmental complementarity and suprasegmental parallelism gradually shift *in concert* to produce this result. The reduced systems stand out against 6- and 7-tone phonetic backgrounds.

In between these systems there is some structural ambivalence. Is the distribution of a single initial consonant enough basis for an increased tone count? Just how much parallelism between potential tonal allophones needs to be stipulated? These questions could be answered variously, but the fact they arise at all is really the issue. Depending on where in the Southern Thai continuum we look, local linguistic facts require specific interpretive clarifications for the question 'how many tones?'.

Furthermore, it may well be possible to develop an operational test to evaluate degree of tone-system 'integralness' at a given point in the continuum. Such a test would have both sociolinguistic and psycholinguistic dimensions, and could provide interesting insight into how tone systems change.

N O T E S

1. The Thai National Research Council has facilitated field research reported here, which was conducted in 1976-77 with support from the Ford Foundation Southeast Asia Fellowship Program. The Southern Thai Cultural Studies Institute, directed by Acan Suthiwong Phongphaibun at Sri Nakharinwirot University, Songkhla, has provided much helpful assistance.
2. White Tai forms in this paper are from Donaldson (1970) cited in a standardised form. For Chuang, cf. Moskalev 1971. For other Tai languages see bibliographies in Gething, Harris and Kullavanijaya (1976).
3. cf. Wang 1967. Figures here are derived from normalised sound spectrogram contours. There is general agreement with earlier reports of Haas (1958), Henderson (1959), etc., but some departure from Haudricourt (1972:82, particularly for his tone classes represented by '*chicken*' and '*father*').
4. This is a Trang variety departing slightly from the schema here; see also following discussion.
5. For an opposite sort of shift, cp. the register common in Southern Thai shadow puppetry for narrative summary: tonal distinctions are submerged in monotone chanting, and segmentals must be relied on to convey the message (perhaps rather predictable anyway).
6. Of course the ñ-/y- distinction is actually a gradient one with varying degrees of nasal air-stream reduction along with lenition of palatal closure. Similarly gradient is the urban elite to rural social transition. For historical discussion of the distinction, cf. Duangduan

Suwatee and Pranee Kullavanijaya (1976; in Gething, Harris and Kullavanijaya, 1976).

7.  $y^*$  seems to be restricted to 20 items, of which 7 form a final-stop subclass to be mentioned later. In addition to items noticed by Egerod (1961:73) and Brown (1965:146) are  $e$ -category items:

$y\text{a}\eta$  'to test depth'  
 $y\text{ɔ}\text{ɔ}\text{n}$  'to slacken'  
 $y\text{ɔ}\text{ɔ}\eta$  'betel tray'

Note also that  $y\text{i}k$  'to pinch' and  $y\text{a}\text{a}\eta$  'rubber tree' are  $y^*$ , but not  $y\text{i}k$  'curly' and  $y\text{a}\text{a}\eta$  'Dipterocarpus tree'; 'pond heron'. Respective pairs are both homographs and homonyms in Central Thai.

8. A few exclamatory forms provide further support, e.g.  $\text{doo}$  (as though  $\text{ad}$ ) 'way over there'; for 7°N rural varieties they are badgering the tonal reduction, as are urban hybrid forms creeping into rural speech. Thus younger rural speakers will accept and produce  $y$ - alternates for  $\tilde{n}$ - (but not  $\tilde{n}$ - alternates for  $y$ - for  $y^*$ ).

BIBLIOGRAPHY

BROWN, J. Marvin

- 1965 *From Ancient Thai to Modern Dialects*. Bangkok: Social Science Association Press of Thailand.

COURT, Christopher

- 1975 'The Segmental and Suprasegmental Representation of Malay Loanwords in Satun Thai: A Description with Historical Remarks'. In: Jimmy G. Harris and James R. Chamberlain, eds *Studies in Tai Linguistics in Honor of William J. Gedney*, 67-88. Bangkok: Central Institute of English Language.

DONALDSON, Jean

- 1970 *Tai-Vietnamese-English Vocabulary*. Saigon: Bo Giao-Duc Xuat Ban. (With D. Nhim.)

EGGEROD, Søren

- 1961 'Studies in Thai Dialectology'. *Acta Orientalia* 26/1-2: 43-91.

GANDOUR, Jack

- 1975 'Counterfeit Tones' in the Speech of Southern Thai Bidialectals. Paper presented at the Annual Meeting of the Linguistic Society of America, December 1975.

GETHING, Thomas W., Jimmy G. HARRIS and Pranee KULLAVANIJAYA, eds

- 1976 *Tai Linguistics in Honor of Fang-Kuei Li*. Bangkok: Chulalongkorn University Press.



HAAS, Mary R.

- 1958 'The Tones of Four Tai Dialects'. *Bulletin of the Institute of History and Philology, Academia Sinica* 29:817-26.

HAUDRICOURT, A.G.

- 1972 'Two-way and Three-way Splitting of Tonal Systems in some Far Eastern Languages'. (Trans. C. Court.) In: Jimmy G. Harris and Richard B. Noss, eds *Tai Phonetics and Phonology*, 58-86. Bangkok: Central Institute of English Language.

HENDERSON, E.J.A.

- 1959 'The Tones of the Tai Dialect of Songkhla'. *Bulletin of the Institute of History and Philology, Academia Sinica* 30:233-5.

JONES, R.B.

- 1965 'On the Reconstruction of Proto-Thai'. *Lingua* 14:194-229.

MOSKALEV, A.A.

- 1971 *Grammatika Yazika Chzhuan*. Moscow: Nauka.

PANUPONG, Vichin

- 1972 'Pronunciation Characteristics of Words in the Songkhla Dialect'. [In Thai.] In: Jimmy G. Harris and Richard B. Noss, eds *Tai Phonetics and Phonology*, 87-104. Bangkok: Central Institute of English Language.

WANG, William S-Y.

- 1967 'Phonological Features of Tone'. *International Journal of American Linguistics* 33/2:93-105.



# TONAL RULES FOR ENGLISH LOANWORDS IN THAI\*

JACK GANDOUR

EDITOR'S NOTE: This paper was first published in *Studies in Tai and Mon-Khmer Phonetics and Phonology in Honour of Eugenie J.A. Henderson*, ed. by T.L. Thongkum, V. Panupong, P. Kullavanijaya, and M.R.K. Tingsabadh, 1979, Bangkok: Chulalongkorn University Press; republished here with the kind permission of Professor Theraphan Luangthongkum, Chulalongkorn University.

## 0. INTRODUCTION

In cases of contact between two languages, we often find that one language will borrow words from the other. The borrowing language normally adapts the loanwords, to a greater or lesser degree, to fit into its phonological system. Depending on the extent to which the modifications the loanwords undergo are phonologically (or phonetically) motivated, we have external evidence in support of the psychological reality of the sound patterns of the borrowing language, as well as evidence that bears on models of speech production and perception. Much of the linguistics research to date on loanword phonology has focussed exclusively on the borrowing of segmental features, and the particular modifications they undergo in the mapping process from the donor language to the borrowing language.

There has been comparatively little research done on the borrowing of suprasegmental features and, in particular, on the borrowing of pitch contours. Egerod (1959) provided tonal correspondences from the Chaochow and Chaozhou dialects of Chinese to Thai. And more recently, Maddieson (1977) discussed Hausa loanwords in Gwari, Nupe and Kpan. In both the Egerod and Maddieson studies, the donor language was a tone language and the borrowing language was a tone language.

Since it is commonly known that pitch is the principal feature underlying the perception of tone as well as one of the most important dimensions (others include length, loudness, vowel quality) underlying

---

\* I wish to particularly thank Sirinapa Jamornmarn for her assistance in collecting data for this paper.

the perception of stress, it is of much theoretical interest to investigate cases of language contact in which the donor language is a stress language and the borrowing language is a tone language, or vice versa. One interesting case in South-east Asia involves contact between the parent language of the Tai language family and Indic, Khmer and Indonesian languages sometime during the 14th century. In this borrowing situation, Proto-Tai was the borrowing language; Indic, Khmer and Indonesian were donor languages. Among Tai historical-comparativists, Proto-Tai is generally considered to be a tone language, having three tonal categories A, B and C associated with those syllables not ending in a stop segment; Indic, Khmer and Indonesian languages at this stage of historical development are all considered to be non-tonal. The manner in which Indic, Khmer and Indonesian loanwords were adapted into Proto-Tai suggest underlying phonological factors determined the tonal representation of syllables in loanwords borrowed from non-tonal languages. Benedict (1942:598), in his investigation of the possible genetic affiliation between Tai and Indonesian languages, commented that it is probably significant that almost all Thai roots that have correspondences in Indonesian languages are classified under the Proto-Tai tonal category A. In his historical-comparative survey of the Tai language family, Gedney (1967) too noted that syllables ending in a sonorant segment in Indic and Khmer loanwords were all assigned Proto-Tai tone A when adapted to the Tai parent language. He goes on to suggest that "tone A was the normal level tone, with tones B and C so markedly different from it as never to be used in pronouncing the syllables of words borrowed from a toneless language."

More recently, Court (1975), in his study of the language contact situation between the Satun dialect of Thai, the tonal borrowing language, and Malay, the non-tonal donor language, also found evidence of phonological factors mediating between the stress and intonational patterns of a non-tonal language, and the tonal representation of these suprasegmental dimensions in a tonal language. In particular, he reported that the high rising-falling tone is assigned to syllables ending in a sonorant segment that occur in phrase-final positions, and that the high level tone is assigned elsewhere. Court suggests that the high rising-falling tone of the Satun Thai dialect is the most suitable tone for representing the falling intonation nucleus of Malay loanwords because of its high starting point, as compared to the two other falling tones in the Satun Thai tonal inventory, both of which have lower starting points.

One other study (Kiu 1977) of a borrowing situation between a tonal language and a non-tonal language involves Cantonese and English. Again, the tonal language, Cantonese, is the borrowing language; the stress language, English, is the donor language. In bisyllabic English loanwords, Kiu reported two different tonal patterns that correlate with two different stress patterns. Bisyllabic loans with stress on the first syllable ( $\acute{SS}$ ) are represented tonally in Cantonese with the high falling tone, or the high level tone, a common sandhi variant in non-phrase final position, on the first syllable, and the high rising tone on the second syllable ( $S^{53}S^{35}$  or  $S^{55}S^{35}$ ); bisyllabic loans with stress on the second syllable are represented tonally in Cantonese with the mid level tone on the first syllable, and the high level tone on the second syllable ( $S^{33}S^{55}$ ). While the second bisyllabic tonal pattern ( $S^{33}S^{55}$ ) correlates nicely with the bisyllabic English stress pattern ( $S\acute{S}$ ), it is hard to account for the first tonal pattern ( $S^{53}S^{35}$  or  $S^{55}S^{35}$ ) in terms of the pitch contours associated with the bisyllabic stressed-unstressed pattern ( $\acute{SS}$ ). The fact that the high rising tone occurs on the second syllable in these loanwords suggests that non-phonetic factors are also at work in converting stress patterns into tonal patterns. Perhaps, the high rising tone serves to indicate that the word is clearly a foreign item in the Cantonese lexicon.

This paper investigates yet another case of language contact between a tone language and a stress language - this time between (Standard) Thai, the national language of Thailand, and English. In this contact situation, Thai is the borrowing language, English the donor language. The aim of the paper is to formulate the rules for converting the stress and intonational patterns of English into Thai tonal categories, and to determine the extent to which the resultant tonal patterns can be accounted for in terms of the pitch contours associated with the English stress patterns. The manner in which the segmental properties of English loanwords are modified will be mentioned only where it is relevant to our discussion of tone and syllable structure. For a more detailed discussion of segmental modifications of English loanwords in Thai, see Henderson (1951).

## 1. THAI TONES/SYLLABLE STRUCTURE

Thai has five contrastive lexical tones: mid ( ), low (``), falling (^), high (') and rising (\*). These tones are illustrated below in the familiar Chao (1930) tone letters.

(1)	naa	↘	32	<u>mid</u> falling	'field'
	na'a	↘	21	<u>low</u> falling	(a nickname)
	na'á	↘	51	high <u>falling</u>	'face'
	na'á	↗	45	<u>high</u> rising	'aunt'
	na'á	↗	24	low <u>rising</u>	'thick'

Constraints on the distribution of these lexical tones depend on syllable structure. All five tones may occur on syllables ending in a long vowel or sonorant segment (CV:, CV:C<sub>f</sub>, CVC<sub>f</sub> where C<sub>f</sub> = m n ŋ w j, hereafter, 'smooth' syllables). On syllables that consist of a long vowel followed by a non-sonorant segment (CV:C<sub>f</sub> where C<sub>f</sub> = p t k, hereafter, 'long checked' syllables), only the low and falling tones are permitted. On syllables that consist of a short vowel followed by a non-sonorant segment (CVC<sub>f</sub> where C<sub>f</sub> = p t k ?, hereafter, 'short checked' syllables), only 'short' variants of the low and high tones are allowed. Of particular relevance for the assignment of tones to English loanwords is that the high tone may not occur on long checked syllables in the native Thai vocabulary.

The constraints on the distribution of tones in English loanwords, while not identical to those that apply to native Thai lexical items, also depend on syllable structure. Each syllable of an English loanword is assigned a tone. On smooth syllables, only the mid and falling tones are permitted; the low, high and rising tones do not occur on smooth syllables in English loanwords. On long checked syllables, the low and high tones occur, but not the falling tone. The occurrence of the high tone on long checked syllables in English loanwords is at variance with native Thai syllable structure constraints on the distribution of tones. As mentioned above, the high tone is prohibited on long checked syllables in native Thai vocabulary (cf. Henderson 1949). On short checked syllables, both the short low and high tones may occur. In addition to these syllable structure constraints, we will see that tonal distribution in English loanwords further depends on the position of the syllable within a word.

## 2. ENGLISH LOANWORDS

### 2.1. MONOSYLLABLES

The assignment of tones to monosyllabic English loanwords depends on syllable structure alone; syllables interpreted as smooth as assigned mid tone, syllables interpreted as checked, short or long, are assigned high tone.

(2)	'cream'	khri:m	'soup'	súp
	'queue'	khiw	'golf'	kóp
	'term'	thə:m	'Fab'	fé:p
	'farm'	fa:m	'chalk'	chók
	'Jew'	jiw	'wig'	wík
	'share'	che:	'fluke'	flúk
	'bar'	ba:	'Lux'	lók
	'beer'	bia	'set (tennis)'	sét
	'fair'	fe:	'shirt'	chát:t
	'film'	fi:m	'cake'	khé:k
	'free'	fri:	'oats'	ʔó:t
	'ream'	ri:m	'oak'	ʔó:k
	'cheer'	chia	'card'	ká:t
	'mile'	maj	'yeast'	jí:t
	'fan (admirer)'	fe:n	'Jeep'	cí:p
	'view'	wiw	'wheat'	wít

Other monosyllabic English loanwords clearly show that tonal assignment is made according to what is perceived as the English syllable structure, and not according to the eventual phonetic realisation of the corresponding Thai syllable.

(3)	'bank'	bén	*bən
	'pump'	pám	*pam
	'tent'	tén	*ten
	'belt'	bén	*ben
	'tank'	thén	*thən

These English monosyllables end in a non-sonorant segment and, according to rule, are assigned high tone. But phonetically, these English loanwords end in a sonorant segment, which might lead us to expect these loanwords to conform to the rules for smooth syllables. If we assume that tonal assignment occurs 'prior to' the simplification of syllable-final consonant clusters, then the English loanwords in (3) are in agreement with the general pattern of tonal distribution in monosyllables. Without this assumption, we are left no alternative but to treat them as exceptions!

Besides this interaction between syllable-final cluster simplification and tonal assignment, we also find an interaction between the distributional constraints on Thai diphthongs and the rules for assigning tones to English monosyllables.

- |     |                  |     |      |
|-----|------------------|-----|------|
| (4) | 'pipe'           | páj | *paj |
|     | '(Mickey) mouse' | máw | *maw |
|     | 'night (club)'   | náj | *naj |

Thai prohibits the diphthongs aj and aw on checked syllables. Again, we see that tones are assigned on the basis of the English syllable structure. Although the English loanwords in (4) exhibit smooth syllable structures, they carry the high tone because the corresponding English syllables end in a non-sonorant segment.

There are also English monosyllables that are restructured in Thai as bisyllabic loanwords, due to restrictions on syllable-initial consonant clusters.

- |     |                       |        |
|-----|-----------------------|--------|
| (5) | 'steak'               | saték  |
|     | 'skirt'               | saká:t |
|     | 'Scotch (tape)'       | sakót  |
|     | 'smart (fashionable)' | samá:t |
|     | 'switch'              | sawít  |
|     | 'spring (season)'     | saprin |
|     | 'swing'               | sawin  |

In particular, Thai does not permit syllable-initial consonant clusters beginning with an alveolar fricative. Thus, the English loanwords in (5) are modified to meet this restriction on syllable onsets. The second syllable of these bisyllabic loanwords follows the rules already established for English monosyllables - mid tone on smooth syllables, high tone on checked syllables. The first syllable in these loanwords carries a mid tone, that results from the application of a tone sandhi rule that neutralises the distinction between high and low tones on CV syllables at regular speech tempo (for detailed formulation of this tone reduction rule, see Gandour 1976).

Finally, there are a few exceptions to the rules for assigning tones to English monosyllabic loanwords.

- |     |                   |        |        |
|-----|-------------------|--------|--------|
| (6) | 'sport'           | sapò:t | sapó:t |
|     | 'Ford'            | fò:t   | fó:t   |
|     | 'suit'            | sù:t   | sú:t   |
|     | 'serve (wait on)' | sà:p   | sá:p   |
|     | 'stamp'           | satəm  | satém  |
|     | 'fuse'            | fiw    | fíw    |



## 2.2. POLYSYLLABLES

Let us next consider bisyllabic English loanwords. Those listed in (7) provide a representative sample (primary stress on English words is marked with an acute accent).

(7)	'vísà'	wi:sâ:	'doctor (Dr)'	dók-tâ:
	'fáshion'	fɛ:chân	'Parker (pen)'	pákkâ:
	'dóllar'	dɔ:lâ:	'Pepsi'	pépsî:
	'pártý (affair)'	pɑ:tî:	'rugby'	rákbî:
	'cócóa'	ko:kô:	'taxi'	théksî:
	'bówling'	bo:lîŋ	'Néstle's'	nés-lê:
	'ríflē'	rajfân	'pízza'	phítsâ:
	'(Coca) Cólà'	kho:lâ:	'tráctor'	thréktâ:
	'ráyon'	re:jô:n	'ápple'	ʔéppân
	'(President) Cárter'	kha:tê:	'lécture'	lékchâ:
	'rádar'	re:dâ:	'níckel'	níkkân
	'Ú.S.O.M.'	ju:sô:m	'swéater'	sawéttâ:
	'(eye) sháadow'	che:dô:	'whiskey'	wítsakî:
	'álbum'	ʔa:labâm		

Except for the second syllable, these bisyllabic loanwords can be accounted for with the rules already established for monosyllables. Instead of a mid tone, as we might expect, based on the evidence from those monosyllables ending in a sonorant segment, the second syllable in these loanwords is assigned a falling tone. The fact that we find a falling tone, as opposed to a rising tone, on the final syllable of these loanwords suggests a natural phonetically-motivated explanation in terms of English and Thai stress patterns. The stressed-unstressed English pattern in (7) correlates with a falling pitch contour. Since Thai rhythm requires that the last syllable in a phrase be stressed, it would appear that the falling pitch contour has been preserved in the Thai pronunciation, but that the point of the fall has been shifted to the final syllable in accordance with Thai rhythmic constraints.

Though this phonetically-motivated account has a lot of appeal, it cannot be the whole story, as evidenced in bisyllabic loanwords whose second syllable ends in a non-sonorant segment.

(8)	'bónus'	bo:nát
	'sándwich'	sɛnwít
	'dónut'	do:nát
	'ténnis'	thennít
	'fóremost'	fɔ:mó:t
	'jácket'	cékkét

Assuming that the Thai tonal patterns reflect the English stress patterns, we would surely expect the low tone on the final syllable of this set of loanwords. But instead what we find is the high tone. Since the English stress pattern in (8) displays an overall falling pitch contour, we cannot explain, from a strictly phonetic point of view, the overall rising pitch contour on the borrowed forms in Thai. There are, however, other bisyllabic loanwords which do seem to support a phonetic interpretation.

(9) 'Égypt'	ʔi:jìp
'Cólgate'	khɔ̀nkè:t
'cóncrete'	khɔ̀nkrì:t
'lípstick'	lípsatìk
'hótdog'	hóttdò:k
'Klé'nex'	khi:nèk
'crédit'	khre:dìt
'téchnique'	théknìk
'(air) hóstess'	hóssatè:t
'plástic'	phléssatìk
'(Robert) Rédford'	rétfð:t
'pássport'	phéssapð:t
'cóntact (lens)'	khɔ̀nthèk
'Éurope'	júʔrò:p
'Píttsburgh (Pa.)'	phíssabè:k

The English words in (9) all display a falling pitch pattern with the primary stress falling on the first syllable. The Thai adaptation of this set of words reveals that the tonal pattern more closely approximates the English stress pattern. In particular, we observe that the checked syllables at the end of these loanwords are all assigned the low tone. This tonal pattern corresponds to the lowered pitch on the second syllable of these bisyllabic English words. When checked syllables occur in positions other than at the end of polysyllabic words, they are always assigned the high tone. Given the English loanwords in (8) and (9), it would appear then that we have competing strategies for the adaptation of English stress patterns into the Thai language. Those loanwords in (9) suggest that the process of adaptation is phonetically motivated, while those in (8) indicate perhaps a conventionalised reading pronunciation of English orthography.

The bisyllabic English loanwords in (10), however, clearly suggest some phonetically-based adaptation strategy.

(10)	'shampoo'	chemphu:	*chemphû:
	'Tŭ'	thi:wí:	*thi:wî:
	'hifí'	hajfaj	*hajfâj
	'hotél'	ho:ten	*ho:tên
	'revuê'	ri:wiw	*ri:wîw
	'psychó(logy)'	sajkho: (vb.)	*sajkhô:

We might expect these bisyllabic loanwords to conform to the tonal pattern established for the English loanwords listed in the left column of (7). But instead of the falling tone being assigned to the second syllable of the loanwords in (10), we observe that the second syllable is pronounced with the mid tone. Unlike the English words in (7), which display primary stress on the first syllable, the words listed in (10) all display primary stress on the second syllable. These English source words have an overall rising stress pattern, the second syllables being comparatively higher in pitch than the first, and longer. If these words were to be adapted with the falling tone of the second syllable, the resultant tonal pattern would be considerably different from the perceived stress pattern. Thus, the final syllables of these bisyllabic words are assigned a mid tone which results in a closer approximation to the English stress pattern. It is surely significant that English syllables ending in a sonorant segment, at the end of polysyllabic words, are normally assigned the falling tone, except for such bisyllabic words in (10), and trisyllabic words and phrases such as 'violin' waj'o:lin, \*waj'o:lîn; 'Oil of Oláy' ʔɔ: ʔáf ʔu:lâ:n, \*ʔɔ: ʔáf ʔu:lâ:n, which all carry primary stress on the final syllable.

Other bisyllabic English loanwords, however, show exceptional behaviour.

(11)	'sóda'	so:da:	*so:dâ:
	'Níxon'	níksǎn	*níksân
	'billiards'	binli'at	*binli'at, *binli'at
	'nécktie'	néktháj	*nékthâj
	'Christmas'	khrítsamâ:t	*khrítsamâ:t, *khrítsamá:t

Let us next turn our attention to English source words consisting of more than two syllables.

(12)	'Alsáccian'	ʔanse:chĩan
	'pollútion'	pho:lu:chân
	'corrúption'	khɔ:rápchân
	'compúter'	khɔmphiwtâ:
	'Míami'	majʔɛ:mî:
	'Chicágo'	chíʔka:kô:

'(North) Carolina'	kha:ro:la jnâ:
'(North) Dakota'	da:ko:tâ:
'Nebraska'	ne:bréssakâ:
'Alaska'	?aléssakâ:
'Oklahoma'	?o:kla:ho:mâ:
'Alabama'	?e:la:ba:mâ:
'Montana'	montha:nâ:
'hamburger'	hembæ:kê:
'badminton'	bétmintân
'Rotary (Club)'	ro:tarî:
'lottery'	lótтарî:
'battery'	béttarî:
'Kennedy'	khennadî:
'hydrogen'	hajdro:cên
'Hungary'	hanka:rî:
'Washington'	wá?chintân
'furniture'	fæ:nicâ:
'technology'	thékno:locî:
'Impérial'	?imphi:riân
'aluminum'	?alu:mi:niâm
'uranium'	jure:niâm
'biology'	baj?o:lo:cî:
'Niagara (Falls)'	naj?eŋkarâ:
'Democrat'	demmo:krèt
'Chevrolet'	chépro:lèt
'Massachusetts'	métsa:chu:sè:t
'chocolate'	chókko:lèt
'Cadillac'	kha:di:làk

The polysyllabic English loanwords in (12) point to a conventionalised reading pronunciation of English orthography. The rules for tonal assignment are based strictly on the interpretation of English syllable structure. Those syllables interpreted as smooth receive the mid tone in non-final position, the falling tone in final position; those syllables interpreted as checked receive the high tone in non-final position, the low tone in final position. Short open syllables in English source words that occur between a primary stressed syllable and a following syllable are assigned a mid tone in accordance with the tone reduction rule in Thai. Since the tonal patterns remain fixed in the adaptation of variable stress patterns found in polysyllabic English words, we cannot attribute the resultant tonal patterns to perceptual

interpretation of the variable pitch contours associated with the English stress patterns.

English loanwords for the names of certain countries also point to a conventionalised reading pronunciation.

(13) 'Índia'	ʔindia	*ʔindíʔa
'Nórway'	nɔːraweː	*nɔːrawêː
'Gérman(y)'	jəːraman	*jəːramân
'Fínland'	finlɛːn	*finlêːn
'África'	ʔéffriːkaː	*ʔéffriːkâː
'Arábia'	ʔaːraːbia	*ʔaːraːbíʔa
'Indonésia'	ʔindoːniːsia	*ʔindoːniːsíʔa
'América'	ʔameːrikaː	*ʔameːrikâː
'Switzerland'	sawítsalɛːn	*sawítsalêːn
'The Nétherlands'	neːthaleːn	*neːthalêːn
'Túrkey'	təːrakɪː	*təːrakîː

The smooth syllables in the loanwords of (13) are assigned the mid tones, instead of the expected falling tone. The fact that the names of cities and countries in the English loanwords of (12) do follow the more general tone pattern, while those in (13) do not, again suggests competing strategies in the adaptation of English stress patterns.

Some other English loanwords which display exceptional behaviour in terms of the general tonal patterns found in adaptation of English stress patterns are listed in (14).

(14) 'óxygen'	ʔóksiceːn
'vítamin'	wíttamin
'Óvaltine'	ʔoːwantin
'vanílla'	wáʔniːlaː
'macaróni'	mákkaroːniː
'báskéttball'	béssakétbɔːn
'mágageine'	mékkasiːn
'Hólllywood (Calif.)'	hɔːliwúːt

### 3. SUMMARY

While both phonetic and non-phonetic factors appear to interact in determining the eventual tonal representation of English stress patterns, the rules for assigning tones, as summarised in (15), do apply to the overwhelming majority of English loanwords in Thai.

(15)

	monosyllabic words	polysyllabic words	
		non-final position	final position
smooth syllable	mid	mid	fall
checked syllable	high	high	low

Some of the exceptions to these rules can be accounted for with phonetic, morphological, and/or semantic information, as shown by the loanwords in (10) and (13), yet other exceptions cannot, as shown by the loanwords in (14). With the increasing influence of English as a second language in Thailand, we will have further opportunity to observe how these tone rules are extended or restricted in application in the process of borrowing words from a stress language.

BIBLIOGRAPHY

BENEDICT, Paul K.

- 1942 'Thai, Kadai, and Indonesian: A New Alignment in South-eastern Asia'. *American Anthropologist* 44:576-601.

CHAO, Y.R.

- 1930 'A System of "Tone Letters"'. *Le Maître Phonétique* 32: 24-7.

COURT, C.

- 1975 'The Segmental and Suprasegmental Representation of Malay Loanwords in Satun Thai: A Description with Historical Remarks'. In: J.G. Harris and J.R. Chamberlain, eds *Studies in Tai Linguistics in Honor of William J. Gedney*, 67-88. Bangkok: Central Institute of English Language, Office of State Universities.

EGEROD, S.

- 1959 'Swatow Loan Words in Siamese'. *Acta Orientalia* 23:137-56.

GANDOUR, J.

- 1976 'A Reanalysis of Some Phonological Rules in Thai'. In: T.W. Gething, J.G. Harris and P. Kullavanijaya, eds *Tai Linguistics in Honor of Fang-Kuei Li*, 47-61. Bangkok: Chulalongkorn University Press.

GEDNEY, W.J.

- 1967 Future Directions in Comparative Tai Linguistics. Manuscript. University of Michigan, Ann Arbor.

## FENDERSON, E.J.A.

- 1949 'Prosodies in Siamese, a Study in Synthesis'. *Asia Major* (New Series) 1:189-215. Reprinted in: F.R. Palmer, ed. *Prosodic Analysis*, 27-53. London: Oxford University Press, 1970.
- 1951 'The Phonology of Loanwords in some South-East Asian Languages'. *Transactions of the Philological Society* 1951:131-58.

## FIU, K.L.

- 1977 'Tonal Rules for English Loan Words in Cantonese'. *Journal of the International Phonetic Association* 7/1: 17-22.

## MADDIESON, I.

- 1977 'Tone Loans: A Question Concerning Tone Spacing and a Method of Answering it'. *UCLA Working Papers in Phonetics* 30:49-83.



# COMPLICATIONS IN TEMPORAL PREVERBS AND THEIR SEMANTIC INTERPRETATION<sup>1</sup>

UDOM WAROTAMASIKKHADIT

## 1. INTRODUCTION

The term 'temporal preverb'<sup>2</sup> is used in this paper, in some sense, overlaps with the more familiar terms 'aspect', 'tense', and 'auxiliary'. The terms 'aspect' and 'tense' which are normally used in the Indo-European languages are mostly concerned with affixing, and the term 'auxiliary', which is not concerned with affixing, covers a wider range than the term 'temporal preverb' in this paper. The author will exclude some other preverbs such as *khon* 'may', *khuan* 'ought to', *ʔàat* 'probably', and *tôn* 'must' which are not temporal preverbs. This analysis posits that all of the Thai temporal preverbs occur preverbally in the deep structure; they, thus, are legitimately called 'preverbs'.

## 2. TEMPORAL PREVERBS AND THEIR SEMANTIC FEATURES

This paper presents three kinds of temporal preverbs.

### 2.1. *cà* AS A PREVERB

Normally *cà* is interpreted as a future particle<sup>3</sup> containing the feature [+future]. Let us consider the following sentences.

- (1) *khăw cà kin khâaw*  
*he future eat rice*  
'He is going to eat.'
- (2) *khăw cà mây kin khâaw*  
*he future not eat rice*  
'He is not going to eat.'

*cà* in (1) and (2) signifies a regular future.<sup>4</sup> A negative may be introduced to the position following *cà* by a negative transformation in the surface structure.

## 2.2. kamlan-yùu AS PREVERBS

This analysis proposes that kamlan<sup>5</sup> and yùu<sup>6</sup> preverbs can occur as syntactic echoes<sup>7</sup> to each other in the deep structure where both or either one of the two may occur in the same simple sentence. The features [+progressive] and [+definite] are assigned to kamlan and the features [+progressive] and [-definite] to yùu. Let us consider the following sentences.

- (3) khǎw kamlan            kin khâaw  
       he    progressive eat rice  
       *'He is (definitely) eating.'*
- (4) khǎw kin khâaw yùu  
       he    eat rice    progressive  
       *'He is (indefinitely) eating.'*
- (5) khǎw kamlan            kin khâaw yùu  
       he    progressive eat rice    progressive  
       *'He is eating.'* (emphatic)

yùu, an indefinite syntactic echo of kamlan, is transformed to the position following a verb phrase in the surface structure. (3) and (4) have nearly the same semantic interpretation. (3) is taken to be more definite than (4) in the speaker's view. Any one of the sentences (3) to (5) can be the answer to any one of the following questions.

- (6) khǎw kamlan            tham ?àray  
       he    progressive do    what  
       *'What is he (definitely) doing?'*
- (7) khǎw tham ?àray yùu  
       he    do    what    progressive  
       *'What is he (indefinitely) doing?'*
- (8) khǎw kamlan            tham ?àray yùu  
       he    progressive do    what    progressive  
       *'What is he (definitely) doing?'* (emphatic)

When the feature [+definite] precedes the feature [-definite] in the same simple sentence, the feature [-definite] becomes [+emphatic]. Thus either (5) or (8) conveys emphaticness in its semantic interpretation.

- (9) [-definite] → [+emphatic] / [+definite] \_\_\_\_
- (10) khǎw kamlan            cà            kin khâaw  
       he    progressive future eat rice  
       *'He will be eating.'* (*'He is about to eat.'*)

- (11) khăw cà        kin khâaw yùu  
       he    future eat rice    progressive  
       *'He will be eating.'*
- (12) khăw kamləŋ        cà        kin khâaw yùu  
       he    progressive future eat rice    progressive  
       *'He will be eating.'* (emphatic)

When cà is preceded by kamləŋ, yùu, or kamləŋ-yùu as in (10), (11), and (12), kamləŋ cà,<sup>8</sup> yùu cà, or kamləŋ-yùu cà will be interpreted as a near future where the action will occur sooner than the one without either kamləŋ or yùu, or both. It can be stated by the following rule.

- (13) [+future] → [+future2]<sup>9</sup> / [+progressive] \_\_\_\_

(11) is taken to be ambiguous with two readings. It either says that he will be (indefinitely) eating in the near future, or (I expect that) he will be (indefinitely) eating. The first reading is considered regular but the second reading must be derived from a different deep structure which will not be treated here.<sup>10</sup>

cà in (12) behaves in the same manner as cà in (10) but its semantic interpretation of emphaticness should be taken into account because both kamləŋ and yùu appear in the same simple sentence as explained in 2.2.

### 2.3. lăaw<sup>11</sup> AS A PREVERB

This analysis will treat lăaw as a temporal preverb preceding kamləŋ-yùu and cà in the deep structure. lăaw is analysed to have the feature [+cessative] and it is later transformed postverbally to the end of the sentence in the surface structure before a performative transformation is applied.<sup>12</sup>

It must be noted here that this analysis will not consider lăaw as the past tense marker as some authors do.<sup>13</sup> The author believes that lăaw has nothing to do with the past tense: it only shows that whatever the action is, it is ceasing and the state or condition of the action is changing. Let us consider the following sentences.

- (14) khăw kin khâaw lăaw  
       he    eat rice    cessative  
       *'He has already eaten.'*
- (15) khăw cà        kin khâaw lăaw  
       he    future eat rice    cessative  
       *'He is going to eat soon.'*
- (16) khăw kamləŋ        cà        kin khâaw lăaw  
       he    progressive future eat rice    cessative  
       *'He will definitely be eating very soon.'*

- (17) khǎw cà      kin khâaw yùu      lǎaw  
       he    future eat rice    progressive cessative  
       *'He will indefinitely be eating very soon.'*

- (18) khǎw kamləŋ      cà      kin khâaw yùu      lǎaw  
       he    progressive future eat rice    progressive cessative  
       *'He will definitely be eating very soon.'* (emphatic)

(14) indicates that the action of eating is ceasing or has ceased. When cà preceded by lǎaw as in (15), the sentence should be interpreted as the future is ceasing; thus, the action will occur in the nearer near future and sooner than the one with kamləŋ cà as in (11), and kamləŋ-yùu cà as in (12) in 2.2. It can be represented by the following rule.

- (19) [+future] → [+future3] / [+cessative] \_\_\_\_

lǎaw kamləŋ cà in (16), lǎaw yùu cà in (17), and lǎaw kamləŋ-yùu cà in (18) indicate that the near future is ceasing and their semantic interpretation becomes the nearest near future which can be expressed by the following rule.

- (20) [+future2] → [+future4] / [+cessative][+progressive] \_\_\_\_

#### 4. CONCLUSION

This analysis of temporal preverbs is obviously different from that of many others who have written about Thai. It aims to reveal the semantic functions of temporal preverbs and their interpretation in sentences which express progressiveness, futurity, and cessativeness. Combinations of any of the three kinds of temporal preverbs lead to different semantic interpretations. Especially when futurity is involved, there are four degrees of the future, the regular future as expressed by cà, the near future as expressed by kamləŋ cà, yùu cà, or kamləŋ-yùu cà, the nearer near future as expressed by lǎaw cà, and the nearest near future as expressed by lǎaw kamləŋ cà, lǎaw yùu cà, and lǎaw kamləŋ-yùu cà. It is hoped that this analysis will be beneficial to linguists and language scholars interested in the Thai language. Further investigation of another use of cà, deriving from a higher sentence in the deep structure, is needed to support this analysis.

N O T E S

1. This is a revised version of a paper presented at the Ninth International Conference on Sino-Tibetan Languages and Linguistics in Copenhagen, Denmark, October 22-24, 1976. I would like to express my sincere appreciation to Søren Egerod and William J. Gedney for their comments which persuaded me to clarify certain points in this paper.
2. The term 'preverb' is used in Anthony, French, and Warotamasikkhadit 1968, but the term 'temporal preverb' is adopted for this paper only. Haas and Subhanka (1948:582) use the term 'adverb-auxiliaries'.
3. See Anthony, French, and Warotamasikkhadit (1968:101).
4. This analysis posits four degrees of the future. *cà* normally conveys the meaning of futurity unless it is derived from a performative sentence or it is in a restricted environment.
5. See Warotamasikkhadit 1976. *kamləŋ* usually precedes a nonstate verb. A state verb like *pen* 'to be' which is derived from *khɪɪ* cannot be preceded by *kamləŋ*.

(i) \**khăw kamləŋ pen phɪi khʂɔŋ phəm*  
*he progressive be brother of I*

(ii) *khăw kamləŋ pen wət*  
*he progressive be cold*  
*'He is having a cold.'*

*pen* in (i) is derived from *khɪɪ* in the deep structure but *pen* in (ii) is an experiential process verb.

6. *yùu*, a syntactic echo of *kamlan*, may convey a slightly similar meaning to the verb *yùu* 'to stay, to remain' in some senses. Noss (1964:182, 186-8) treats *yùu* as a frequent echo of *yan* 'still' and he also defines *yùu* as 'remaining on the scene, unchanged, action continuing, temporary, without necessary future significance'. He correctly states that the sentence containing the preverb *yùu* signifies temporariness. This analysis also realises that temporariness plays a certain role in *yùu* as well as in *kamlan*. The feature [+temporariness] will not help in distinguishing *yùu* from *kamlan*; thus, the feature [+temporariness] is not employed here. Instead this analysis emphasises the definiteness versus the indefiniteness of the action in the speaker's view.

7. Formerly the term 'doublet' was used, but it may mislead some readers to understand that *kamlan* and *yùu* are philologically derived from the same original.

3. Noss (1964:137) gives the meaning of *kamlan cà* as 'about to' and so do Haas and Subhanka (1948:647). *kamlan cà*, *yùu cà*, and *kamlan-yùu cà* will be compared to *lǎaw cà* in 2.3.

9. The numeral following the feature [+future] signifies the degree of the future approaching the present time. The greater the numeral is, the closer to the present time the action will be. The numeral used here starts with 2.

10. *cà* of the second reading can also precede *kamlan* as follows:

- (1) (mǎa khun tǎn) khǎw cà kamlan kin khǎaw yùu  
 when you wake up he future progressive eat rice progressive  
 '(When you wake up), (I expect that) he will definitely be eating.' (emphatic)

(12) can also have *cà* preceding and following *kamlan* as follows:

- (11) khǎw cà kamlan cà kin khǎaw yùu  
 he future progressive future eat rice progressive  
 '(I expect that) he will definitely be eating.' (emphatic)

*cà* in the above sentences clearly indicates a different semantic interpretation from a regular future as in 2.1. The performative transformation of the higher level sentence either superimposes the feature [+determination] on *cà* following *lǎaw kamlan-yùu* or the feature [+determination] in the performative sentences becomes *cà* preceding *lǎaw kamlan-yùu* as expressed by the following rules:

- (iv)  $\left[ \begin{array}{c} [+future] \\ [+determination] \end{array} \right] \rightarrow \text{cà} / \text{---} (l\acute{a}aw)(kamlaŋ-yùu)$

11. At first glance anyone who knows Thai may conclude that láaw and dâay are syntactic echoes as kamlāṅ and yùu in 2.2. dâay can occur before a verb in the surface structure and many authors seem to take dâay as the past tense marker. Panupong and Dhanvarjara (1975:370) state that dâay should be semantically interpreted as the past tense but Haas and Subhanka (1948:403) warn their students not to use dâay in all sentences which have the past tense in English.

This analysis will consider *dâay* as a verb meaning '*to obtain, to get*' rather than a preverb. It requires either a noun phrase complement or a complement noun following it to complete the sentence and it can be negated as all other verbs by taking *mây* in front of it. The temporal preverbs cannot be preceded by *mây*. Let us consider the following sentences.

- (1) khăw dâay kin khâaw  
 he get eat rice  
 'He gets to eat.'
- (11) khăw mây dâay kròot khun  
 he not get angry you  
 'He was not angry with you.'

Actually the translation of *dây* as the past tense is not far-fetched but it does not always signify the past tense in English as in (i). *dây* frequently occur with negative as *mây dây* as in (ii) but it can also occur without *mây* as in (i). This analysis will consider *dây* neither as a syntactic echo of *lây* nor a temporal preverb; it is merely a verb.

12. See Warotamasikkhadit (1975:342-54).

13. Noss (1964:121) treats *lǎaw* as a specific adjective meaning 'already'. Hzas and Subhanka (1948:144) also give the meaning of *lǎaw* as 'already', and so does Panupong (1970:20-1). Panupong and Dhanvarjara (1975:365) refer to *lǎaw* as the past tense.

BIBLIOGRAPHY

- ANTHONY, Edward M., Deborah P. FRENCH and Udom WAROTAMASIKKHADIT  
 1968 *Foundations of Thai, Part 1 and Part 2.* Ann Arbor,  
 Michigan: The University of Michigan Press.
- HAAS, Mary R., and Heng R. SUBHANKA  
 1948 *Spoken Thai.* 2 vols. New York: Henry Holt.
- MOSS, Richard B.  
 1964 *Thai Reference Grammar.* Washington, D.C.: Foreign Service  
 Institute.
- PANUPONG, Vichin  
 1970 *Inter-sentence Relations in Modern Conversational Thai.*  
 Bangkok: The Siam Society.
- PANUPONG, Vichin and Ratri DHANVARJARA  
 1975 'Characteristics of Grammatical Words in the Thai Language'.  
 In: Jimmy G. Harris and James R. Chamberlain, eds *Studies  
 in Tai Linguistics in Honor of William J. Gedney*, 355-76.  
 Bangkok: Central Institute of English Language.
- WAROTAMASIKKHADIT, Udom  
 1972 *Thai Syntax: An Outline.* The Hague: Mouton.
- 1975 'Dependency of Underlying Structure and Final Particles in  
 Thai'. In: Jimmy G. Harris and James R. Chamberlain, eds  
*Studies in Tai Linguistics in Honor of William J. Gedney*,  
 342-54. Bangkok: Central Institute of English Language.



- 1976      'Peculiarities of the Thai Substantive Verb pen'. In:  
Thomas G. Gething, Jimmy G. Harris and Pranee Kullavanijaya,  
eds *Tai Linguistics in Honor of Fang-kuei Li*, 233-41.  
Bangkok: Chulalongkorn University Press.



# LES PRONOMS PERSONNELS DU KHMER: ORIGINE ET EVOLUTION

SAVEROS POU

## ABRÉVIATIONS

fam.:	familier
IB:	Institut Bouddhique de Phnompenh
IC:	<i>Inscriptions du Cambodge</i> , publiées par G. Coedès
IMA:	Inscriptions modernes d'Angkor
khm.mod.:	khmer moderne
khm.moy.:	khmer moyen
lit.:	littéral(ement)
p.:	personne, personnel
pl.:	pluriel
pr.:	pronom
R.I.:	Rāmakerti I
sg.:	singulier
sk.:	sanskrit
VK:	<i>Vacanānukram Khmaer</i> (1967, 1968)
vx.khm.:	vieux khmer

## 1. PROBLÈME

Un étranger qui s'est donné pour tâche d'apprendre la langue khmère moderne,<sup>1</sup> et qui veut saisir tous les aspects de la langue parlée, ne manque pas d'être dérouté par les difficultés d'emploi des pronoms personnels.<sup>2</sup> Il apprend au départ que, pour parler de lui-même, il doit employer le pronom /kɲom/ en règle générale, l'expression /kɲom preah kərunaa/ ou /kɲom kənaa/,<sup>3</sup> s'il s'adresse respectivement au roi ou à un moine bouddhiste. Puis, il constate que ce dernier, en parlant de lui-même, emploie plutôt le terme ātmābhāb /ʔatmaaphiəp/, ou /ʔatmaa/

tout court. Ayant appris cela, il va s'entretenir avec des amis qui, dans leur parler courant, dénué de conformisme, emploient à la place de /knom/ des termes *a priori* aberrants,<sup>4</sup> tels que ge /kɛɛ/ (pr.p., 3<sup>e</sup> pl.), yoeñ /yœɛ/ (pr.p., 1<sup>ère</sup> pl.), eñ /'æɛ/ (pr.p., 2<sup>e</sup> sg.), gnā /kniə/ (pr.p., 3<sup>e</sup> sg. ou pl.; morphème exprimant la voix réciproque des verbes d'action).

Si l'étranger veut parfaire ses connaissances en matière de pronoms personnels, il devra ajouter au répertoire des termes usuels toute une liste de termes d'adresse ou 'appellatifs', dont l'usage est fort recommandé à la 2<sup>e</sup> et la 3<sup>e</sup> personnes du singulier.<sup>5</sup>

Si maintenant l'étranger se place du point de vue de la linguistique pour étudier le problème des pronoms, il s'aperçoit que le khmer est très pauvre en pronoms personnels. Par exemple, dans une excellente étude de H.J. Pinnow (1965), un simple coup d'oeil permet déjà de constater que le khmer n'est pas plus riche en pronoms personnels que les autres langues môn-khmères. On relève añ /aŋ/ (1<sup>ère</sup> sg.), yoeñ /yœɛ/ (1<sup>ère</sup> pl.), eñ /'æɛ/ (2<sup>e</sup> sg.), vā /wiə/ (3<sup>e</sup> sg. et pl.), ge /kɛɛ/ (3<sup>e</sup> pl.), gāt' /kəɛt/ (3<sup>e</sup> sg. et pl.).<sup>6</sup>

Si l'on prend le terme de 'pronoms personnels' au sens le plus étroit, c'est-à-dire si l'on entend par là les termes de grammaire servant à se référer aux êtres ou aux objets, on devra alors retenir de la liste ci-dessus, déjà réduite, seulement deux termes: añ et yoeñ. En effet, ces derniers - qui sont d'origine môn-khmère - peuvent être considérés comme de véritables pronoms, car ils sont usités depuis toujours comme tels, exclusivement, et sont bien attestés dans les inscriptions en vieux-khmer (añ et yeñ). Parmi les autres, ge et gāt' ont été des noms communs à l'origine - termes notionnels - tandis que vā était un appellatif d'esclaves de sexe masculin (cf. 2.2.3.2.).

Ces constatations vont nous permettre d'exposer les deux aspects du problème que nous avons choisi pour la présente étude:

1) Les pronoms personnels du khmer moderne: multiplicité des formes, ambivalence de certains termes employés.<sup>7</sup>

2) Origine des termes employés comme pronoms personnels et leur évolution fonctionnelle.

Le second va jeter des lumières sur le premier (formes et fonctions actuelles des pronoms): c'est le principal but visé dans cette étude. Celle-ci s'appuiera sur les textes, qui sont parfois fragmentaires. Aussi, pour quelques rares termes sur lesquels je n'ai pas pu recueillir de renseignements suffisants, je me bornerai à émettre simplement des hypothèses sur leur étymologie.

## 2. LES TYPES

La première tâche qui s'impose est une tentative de classification des pronoms. On pourrait envisager cette entreprise sous plusieurs angles.

2.1. D'après la forme des mots, on peut distinguer des pronoms simples et des pronoms composés.

2.1.1. Parmi les mots simples, citons: khñuṃ 'je', anak 'vous', gāt' 'il, elle', ge 'ils, on ...', yoeñ 'nous'. Comme il est dit plus haut, les appellatifs familiers sont fort usités; en particulier, les termes de parenté (cf. 2.2.3.), qui sont des noyaux de polysémie, peuvent être étendus à tout le groupe.

2.1.2. Parmi les mots composés, il faut distinguer:

2.1.2.1. Les composés formés de deux termes qui doivent leur emploi à un besoin de préciser la position du locuteur par rapport à son interlocuteur. Exemple: khñuṃ (=serviteur) 'je' > khñuṃ pād (=serviteur de vos pieds), 'je' employé par un homme s'adressant respectueusement à une personne qui lui est supérieure;<sup>8</sup> khñuṃ mcās' (=servante du maître), 'je' employé par une femme s'adressant à un(e) membre de la famille royale. Hañ eñ (=toi-toi-même), 'tu, toi' de mépris réservé à une femme de basse condition.<sup>9</sup> Lok grū (seigneur-professeur), 'il, vous' en parlant de, ou à, un moine, un professeur ou un médecin.

2.1.2.2. Les expressions périphrastiques dues à un véritable souci de souligner le contraste de conditions du locuteur et de son interlocuteur, ou de celui qui est l'objet de son discours. Le roi, par exemple, est appelé braḥ karuṇā /preah kərunaa ~ preah koonaa/ 'la Sainte Compassion' (cf. 1.). Par conséquent, celui qui s'entretient avec le roi se réfère à lui-même comme khñuṃ braḥ karuṇā (=esclave ou serviteur de la Sainte Compassion) 'je, moi'.<sup>10</sup> De même, en parlant du Buddha, on dit braḥ aṅg (=le saint corps) 'il';<sup>11</sup> et on s'adresse à un grand dignitaire en disant braḥ tej braḥ guṇ /preah dac preah kun/ (=pouvoir ardent-bienfait) 'vous'. On en déduit aisément que, dans les deux cas, 'je' s'exprime respectivement par khñuṃ braḥ aṅg et khñuṃ braḥ tej braḥ guṇ.

2.2. D'après leur origine, on peut classer les pronoms personnels usuels en trois groupes.

2.2.1. Les pronoms personnels d'origine sont au nombre de deux:

2.2.1.1. Añ 'je, moi'. Terme commun à beaucoup de parlers môn-khmers,<sup>12</sup> il a pris en khmer moderne une nuance dépréciative, et ne doit s'employer que si l'on s'adresse à soi-même - cas de monologue - ou à des interlocuteurs envers lesquels on n'a guère de considération.<sup>13</sup>

2.2.1.2. Yoeñ 'nous', attesté en vx.khm. sous la forme de yeñ, est également commun à un certain nombre de parlers môn-khmers.<sup>14</sup>

2.2.2. D'anciens noms communs, pris dans le vocabulaire khmer, ou empruntés à d'autres langues, ont servi à développer la liste des pronoms: ce qui, d'après les textes, semble répondre à un besoin accru de nuancer la pensée. Ce fait n'est pas sans rapport avec l'évolution de l'organisation sociale. Les Khmers semblent soucieux de pousser très loin l'analyse des groupements humains, de leurs rapports entre eux, et celle des rapports des individus entre eux dans chaque groupe. Le statut d'un individu est défini par le groupe social auquel il appartient: il est prince (stec), roturier (rāstr < sk. rāṣṭra) ou religieux (sañgh < p. saṅgha). A cette distinction sociale primaire, qui cependant n'a pas l'immuabilité de celle des castes, s'ajoute une autre, basée sur l'âge de l'individu, ses qualités intellectuelles ou professionnelles, ses liens de sang avec les membres du groupe. Bref, autant de critères, de valeurs enchevêtrées, marqués par la mutabilité, car constamment remaniés selon l'infinie variété des situations où se trouve placé l'individu à chaque étape de son existence.

Voici un exemple typique tiré du khmer moderne, constitué par deux phrases exprimant exactement la même idée, mais énoncées par deux locuteurs de qualité différente, et se rapportant à deux objets-individus de qualité différente également. Schéma:

Je - regarde - lui - s'en aller

(a) khñuṃ moel lok nimant dau (Le locuteur est un Khmer du commun parlant d'un moine)

(b) ātmā moel gāt' ʈoer dau (Le locuteur est un moine parlant d'un laïc roturier)

Donc, voici les deux couples de:

	(a)	(b)
- agent	khñuṃ	ātmā
- objet	lok	gāt'

C'est peut-être cette conception compliquée de la valeur de l'individu qui explique le choix des lexèmes comme pronoms. Il importe, en effet, de souligner qu'ils n'ont pas été pris au hasard.

Avant tout, le choix était motivé par l'idée de '*la personne, le corps, le soi*'. Vx.khm. khlvan '*le corps*' > mod. '*le corps*' > tu, il (fam.); sk. ātman '*le soi*' > khm.moy. et mod. ātmā '*je, moi*' pour les religieux (cf. aussi 2.1.2.: braḥ aṅg et braḥ tanū).

D'autres termes sont d'anciens noms désignant des groupes humains, qui ont pris une valeur partitive. Vx.khm. anak '*les gens*' > khm.mod. anak /neak/ '*vous, Madame*'. Sk. et vx.khm. jagat '*le monde*' > khm. gāt /kɔət/ '*il, elle*'. Sk. et vx.khm. loka '*le monde*' > khm. lok '*Monsieur, il, elle* (pour des personnes de rang élevé)'. Sk. et vx.khm. gaṇa '*groupe*' > khm. gnā /knɛə/ '*je, il ou elle* (fam.)'.

2.2.3. De nombreux appellatifs peuvent être usités couramment comme pronoms. L'exemple type est vx.khm. vā, appellatif d'esclaves mâles dès le khmer pré-angkorien, qui a donné le pronom khm.mod. vā /wiə/,<sup>15</sup> couramment usité à propos des animaux, dépréciatif pour les êtres, et déictique familial. Et si on se souvient que les appellatifs sont, pour la plupart, des termes de parenté, on se rendra compte de l'ampleur de leur usage, et de leur utilité dans la langue courante. De nos jours, les plus usités comme pronoms couvrant les trois personnes (moi, lui et vous) sont: tā /taa/ '*grand-père*', yāy /yɛy/ '*grand-mère*', bū /puu/ '*oncle*', mīn /miɪŋ/ '*tante*', pañ /bɔɔŋ/ '*aîné*', p-ūn /pʰoon/ ou simplement ūn '*cadet*'. En outre, si on veut montrer plus de déférence à l'égard de l'interlocuteur ou de l'individu-objet, on fera précéder ces termes de l'honorifique lok- (cf. 3.7.).

2.2.3.1. Un exemple du vieux khmer mérite d'être cité en guise d'illustration. Jṭ ou ajṭ signifiait '*ancêtre: grand-père, grand-mère, arrière-grand-père, ...*'. Il se délexicalise en khmer moyen pour devenir un préfixe dénotant '*la vénérabilité*', donc appliqué aux termes désignant les ascendants. En même temps, il sert d'appellatif, et de pronom, appliqué aux garçons et jeunes gens. Jṭ est tombé en désuétude, dans le khmer moderne standard.

2.2.3.2. Un individu, locuteur, interlocuteur ou objet d'un discours, doit être considéré dans son contexte social. Un choix préalable doit être fait: il est l'égal des autres, ou bien il leur est '*supérieur*' ou '*inférieur*'. L'inégalité est définie par la naissance ou le rang social - critère variable - mais surtout par l'âge et les qualités

intellectuelles de l'individu. Ce dernier apprend très tôt à reconnaître toutes les données se rapportant aux autres sujets, et à les doser d'une manière rapide, subtile et efficace pour se situer lui-même dans le groupe. Cet apprentissage, théorique (école, famille), ou empirique, fait partie de l'éducation de tout Cambodgien.

Il est probable, toujours d'après les textes, que ce raffinement soit tardif, car on n'en trouve pas trace dans les inscriptions en vieux khmer. Les textes en khmer moyen décèlent l'élaboration progressive d'un système de pronoms, qui est à peu près stabilisé au XX<sup>e</sup> siècle.

Les Cpāp' recommandent avec insistance l'apprentissage des bonnes manières et d'un langage poli. Et nous avons vu plus haut (cf. 2.2.1.1.) que Cpāp' prus condamnait explicitement l'usage de añ. Il est restreint par l'introduction du nom commun khñuṃ 'esclave > je', devenu bientôt obligatoire pour marquer l'effacement volontaire du locuteur devant les autres sujets.

Ainsi, dans Rāmakerti I, Rām dit toujours añ en parlant de lui-même,<sup>16</sup> tandis que Sītā fait usage du terme khñuṃ '(votre) servante'.

Dans IMA 2 (fin XVI<sup>e</sup> s.), la reine Māhakalyāṇavatī Srīśujātā, en s'adressant au Buddha, dit d'elle-même khñuṃ aṃcass 'esclave du maître'.<sup>17</sup>

Dans IMA 6 datée de 1521 śaka, l'auteur-donateur Ak-hlūṅ Abhayarāj s'adresse courtoisement aux lecteurs en employant le pronom lère sg. khñuṃ. Mais, à l'endroit de l'imprécation, il adresse le vœu de vérité à lui-même; dès lors, il emploie le pronom añ:

doḥh kūn añ-eñ gtīy kmūy añ gtīy doḥh cauṃ añ ... (1.17, 18)  
*'qu'il s'agisse de mes propres enfants, mes neveux ou mes  
 petits-enfants'.* (cf. Lewitz 1971:111)

Ainsi en khmer moyen, certains lexèmes se vidaient petit à petit de leur valeur sémantique originelle pour devenir de simples pronoms. Traités de diverses façons pendant cette période-là, ils ont abouti à constituer un système de pronoms encore en usage à l'heure actuelle.

### 2.3. PERSONNE, GENRE, NOMBRE

Pour classer les pronoms, les linguistes qui étudient les langues de l'Asie du Sud-est se servent des cadres auxquels ils ont été habitués: ceux de la classification indo-européenne basée sur les trois personnes, chacune comportant un singulier et un pluriel, réparties en deux classes (sujet et objet). Si ce cadre est gardé dans le tableau final, c'est parce qu'il est pratique, faute de mieux. En réalité, il est imparfait pour diverses raisons.



2.3.1. Certains pronoms sont ambivalents dans leur référence. Un lexème signifiant 'personne, individu', en devenant pronom, pouvait s'appliquer aussi bien à 'moi' qu'à 'toi' ou à 'lui'. La langue écrite, à tendance puriste, essaie de bien distinguer les usages. Mais la langue parlée, en faisant fi des conventions, n'hésite pas à donner une double valeur à certains pronoms. Ainsi:

ge	'je, moi'	18	'il, elle, ils, on'
gnā	'je, moi'		'il, elle, lui'
eñ	'je, moi'	'tu, toi'	
gāt'		'tu, toi'	'il, elle, lui'

2.3.2. La distinction singulier/pluriel a existé en vieux khmer, lequel opposait añ 'je' à yeñ 'nous'. Mais plus tard, la grammaire a encore cédé le pas à des nuances sémantiques. Le pronom khñuṃ actuel, substitué à añ, ne peut être considéré véritablement comme le singulier de yoeñ. Khñuṃ est un pronom courtois lère sg. ou pl., tandis que yoeñ 'nous' s'emploie en général dans les occasions où les formes sont neutralisées. Pour être polies vis-à-vis d'un interlocuteur, surtout d'un 'supérieur', les personnes d'un groupe doivent dire yoeñ-khñuṃ, ou khñuṃ tout simplement.

La règle générale veut qu'on oppose gāt', 3è sg., à ge, 3è pl., 'il/ils'. Mais, dans la pratique, la valeur sémantique originelle commune, 'les gens', subsiste dans ces deux pronoms. A la distinction singulier/pluriel s'est substituée celle de défini/indéfini. Gāt' remplace certaines personnes en particulier, 'il' ou 'ils', tandis que ge désigne 'les gens en général', d'où la traduction par 'on'.

2.3.3. La distinction de genre n'a jamais existé en khmer. Toutefois, elle se remarque en khmer moyen et moderne dans le cas:

2.3.3.1. Des appellatifs-pronoms 2è p.:

masc.	fém.
ā	me
pā, nāy	nān <sup>19</sup>

2.3.3.2. Des vocables spéciaux, tels que khñuṃ braḥ karuṇā, masc./khñuṃ mcās', fém. (lère p., lorsqu'on s'adresse à des princes). Il est évident que ātmā 'je' ne peut être que masculin, usité par les moines, et on sait que la communauté de nonnes est inexistante.

Il n'est donc pas nécessaire d'insister davantage sur le caractère complexe de ces pronoms personnels et les nuances délicates de leur emploi. Seule une pratique longue et assidue du khmer au sein de la communauté khmère permettra à un étranger de saisir toutes les nuances de la pensée derrière la multitude et la polyvalence des signes.

### 3. ORIGINE

L'examen diachronique des pronoms les plus importants relevés des textes va éclairer le problème.

3.1. Khñuṃ /knom/, nom en khm.mod. '*domestique, serviteur*', pr. lère p. '*je, nous*'. Il était connu seulement comme nom en vieux khmer, kñuṃ ~ khñuṃ, désignant '*les esclaves*'.

J'ai eu l'occasion d'examiner ce mot en détail (cf. Lewitz 1974:167-70) et d'en déterminer l'étymologie. En mettant en évidence le mot de base /num/ '*être jeune*', qui a des congénères dans d'autres langues apparentées, j'ai proposé de voir une corrélation entre deux notions '*être jeune*' et '*être inférieur*' en proto-khmer, comme cela s'est passé dans quelques langues européennes. De la sorte, on s'explique le sens de '*esclave*' de vx.khm. kñuṃ, qui est resté en vigueur aussi longtemps qu'a duré l'esclavage lui-même. Cpāp' Kerti kāl parle des khñuṃ '*esclaves*' et des kaṃṭar '*clients*' de l'entourage d'un notable. (st.4a).

Mais dès le XVI<sup>e</sup> siècle, il commença à se substituer à añ comme pr. lère p. dans la langue courtoise, fait commun à l'Asie du Sud-est, pour devenir actuellement un pr. lère p. courant. Il convient de signaler la forme transitoire añ-khñuṃ ('*moi ou nous-esclaves*'), très courante en khmer moyen, mais guère usitée à l'heure actuelle.

3.2. Ge /keḡ/ a toujours existé en khmer, immuable dans sa forme.

3.2.1. En vieux khmer, il se définit par '*les gens*', à sens collectif. Il était employé:

- Comme nom commun, '*les gens, les personnes*'. Ex: jmaḡ ge (K.74, IC VI, 1.3, VII<sup>e</sup>s.) '*les noms de (ces) personnes*'.

- Comme pronom collectif dans des phrases à corrélatifs '*ceux qui ... ceux là*'. Ex: ge ta sak gi ... ge lanlyañ kaṃluṃ niraya (K.341, IC VI, piedroit Nord, 1.11, VIII<sup>e</sup> s.), '*ceux qui déroberont ... ceux-là tomberont en enfer*'.

- Comme préfixe collectif humain. Exemples tirés de deux inscriptions pré-angkorienues: ta man ge pu caḡ añ oy ta vraḡ (K.49, IC VI, 1.12-13),

'ce que les pu caḥ aṅ offrent au dieu'; ge vraḥ phoṅ tel ... (K.38, IC II,1.8), 'tous les dieux qui ...'.

Un cas particulier, cependant, est celui du composé ge kloṅ que G. Coedès a bien identifié comme nom désignant 'une princesse' (IC IV, p.57). Bien que son commentaire ne soit pas tout à fait exact, je ne le reprends pas maintenant, pour ne pas m'écarter trop du sujet. J'ajouterais seulement que ge kloṅ pré-angkorien désigne, de surcroît, une femme mariée: cela explique probablement la survivance de ge dans certains parlers régionaux comme appellatif d'épouse et pronom (cf. 2.3.1.).

3.2.2. En khmer moyen, le mot ge:

3.2.2.1. Fonctionne comme nom. Ex:

metrī sabb ge

dāṃṇ dvīp ṭa puon lokā (R.I., st.205)

'(Il) ... a de l'amitié pour tous les ÊTRES des quatre continents'.

3.2.2.2. Prend la valeur d'un pronom indéfini pluriel 'ils, les autres, autrui, on'. Ex:

mān kāl trūv ge

lhae e ātmā (Cpāp' Kūn cau, st.23)

'... parfois, (la roue du monde) touche les AUTRES ÊTRES et nous laisse du répit'.

3.2.2.3. Dans certains cas, perd le sens collectif pour s'appliquer seulement à une personne, 'il, elle'. Ex:

ge it dos muoy mūt phaṅ (R.I., st.4299)

'Il (=Bālī) n'a pas commis la moindre faute'.

Ce dernier trait explique l'ambivalence du mot ge dans la langue familière actuelle. Puisqu'il se rapporte aussi à une seule personne, il peut désigner aussi bien 'lui' que 'moi'. D'où en khmer moderne ge 'je moi; ils, les autres, on, eux'.

3.3. Gnā /kniə/ vient du sk. gaṇa.

3.3.1. Attesté en vieux khmer sous plusieurs formes, gaṇa ~ gana ~ ganā ~ gnā, il avait seulement la valeur d'un nom désignant 'le groupe, le groupe de gens'. Ex:

ganā nu kulasantāna vāp Īśānaśiva (K.343, IC VI, piédroit Sud,  
1.4-5, Xe s.)

*'l'entourage et la famille du vāp Īśānaśiva'.*

On le rencontre souvent dans l'expression saṃ gaṇa nu 's'associer à' (lit. être en groupe avec).

3.3.2. En khmer moyen, cette valeur se dédouble par une grammaticalisation partielle du lexème, qui se développe de plus en plus jusqu'en moderne.

3.3.2.1. Gnā, lexème, désigne un 'groupe' d'individus liés par le sang, la profession ou quelque autre intérêt; c'est donc aussi 'le parti'. Un homme sans gnā à l'époque moyenne serait comparable à un loup solitaire. Que dire des autres êtres (animal ou dieu)? Rāmakerti I nous en fournit un exemple:

as' debatā phaṇ sīṇ ṭa ar

par gnā vīvar

rat' cuḥ bT traitrīṇsā (st.46)

*'Tous les dieux, réjouis, conduisent leurs troupes (par gnā) en tumulte, et se précipitent du ciel des Trente-trois'.*

3.3.2.2. Cette notion bien définie de 'groupe' explique le premier stade de la grammaticalisation. On relève cet énoncé dans le Kūn cau:

cūr ceḥ rāp' rak

srakar gnī-gnā (st.27d et e)

*'... apprenez à entretenir de bonnes relations (avec tous les membres de votre famille) et rester en groupe avec eux'.*

Gnā est une particule verbale (gnī n'est qu'un élément de redoublement) exprimant la voix réciproque de srakar 'être avec, coaliser'.

3.3.2.3. Tardivement s'ajoute la fonction pronominale: *le groupe* > *ces gens-là* > *ils*'.

3.3.3. Le khmer moderne a conservé toutes ces fonctions de gnā, aux quelles la langue parlée familière ajoute la valeur pronominale partitive: *'ils, eux* > *lui, il* > *moi, je*'.

3.4. Gāt' /kəət/ est la réduction du sk. jagat.

3.4.1. En vieux khmer, l'emprunt sk. jagat est un nom, ayant à la fois le sens collectif de 'monde', et le sens partitif d' 'individus'. Le

premier sens est le mieux illustré par le nom métaphorique des dieux, *vrah kamrateñ jagat* ou '*seigneur du monde*'. Le deuxième peut être relevé dans l'inscription K.277: *nau ru jagat ta varddhe camnām kalpanā neḥ ...* (piédroit Sud, 1.14, XI<sup>e</sup> s., IC IV), '*Quant aux personnes qui feront prospérer cette fondation ...*'.

3.4.2. En khmer moyen (cf. IMA), on retrouve ce mot réduit en monosyllabe (cf. Lewitz 1968:155-67), *gat* ou *gāt*, et employé comme pronom personnel.

IMA 16a donne: *ta kāl gāt niñ is ayūs* (1.11) (Lewitz 1972a:236), '*au moment où elle allait mourir*'.

IMA 18 contient à 1.21 (Lewitz 1973:171): *doh kūṇ gat gtīy doh kmūy gat gtīy* '*qu'il s'agisse de ses enfants ou de ses neveux*' (*kūṇ gat* = '*enfant d'elle*').

3.4.3. Le mot *jagat* n'est jamais restitué en khmer moderne, tandis que le monosyllabe *gāt* /*kəət*/ subsiste comme pronom 3<sup>e</sup> p. courant au singulier, mais plus rare au pluriel. En outre, la langue parlée l'utilise aussi à la 2<sup>e</sup> p.sg., pour apostropher familièrement un interlocuteur, dont voici un exemple:

*Nae gāt!* *Dhvoe i thñai neḥ?*  
'*Alors, mon vieux! Que fait-on aujourd'hui?*'

3.5. *Anak* est commun aux langues môn-khmères et indonésiennes.

3.5.1. Emprunté à l'indonésien, il apparaît en vieux khmer dès les premières inscriptions, sous la forme de *anak* ou *nak*. Ayant la même valeur que *ge*, il désigne avant tout '*les individus, les personnes*', ou '*on*'. Par exemple: *anak kantai* (K.129, IC II, 1.5) '*les femmes*' (lit. personnes-féminin); *anak si*<sup>20</sup> (K.444, IC II, 1.10) '*les hommes*' (lit. personnes-masculin). Voici un énoncé plus élaboré: *anak ta tut padaḥ anak tut vray oy viṣa ta anak* (1.13, BCAI 1911:205), '*les personnes qui incendient les maisons des (autres) personnes, incendient les forêts, donnent du poison aux (autres) personnes ...*'.

Il a également la valeur d'un agent dans des composés tels que *nakk paṃmre* '*les serviteurs*' (=ceux qui servent), *anak sre* '*les travailleurs des rizières*'.

Il sert enfin de préfixe d'appellatif, comme dans *anak khloñ*, que je propose de comprendre '*femme mariée, épouse*' (cf. 3.2.1.: *ge kloñ*).

3.5.2. Anak garde les mêmes fonctions en khmer moyen, avec cette différence que la fonction d'agent se développe considérablement aux dépens des autres, en s'adjoignant celle de pronom. Voici quelques exemples tirés de Rāmakerti I:

ṭaṇ noḥ braḥ nārāy rāmā

draṇ' braḥ krodhā

pT ṭūc anak ṭut nai aggT (st.1216)

'A ce moment-là, le Seigneur Nārāy-Rāmā entre en *colère*, comme si quelqu'un (=une personne) l'avait brûlé avec du feu'.<sup>21</sup>

ksin noḥ braḥ ṛsT munt

anak būt pāy pT

baṃnūt ka huc hanumān (st.2670)

'Aussitôt, l'ascète, (il) fait trois boules de riz qu'il tend à Hanumān' (anak = il = l'ascète)

Dans Cpāp' Kram, on lit:

kum soec tiḥ ṭīel

anak jā pādhyāy

anak hoṇ smoe mtāy

paṅkoet ktT cpāp' (st.16)

'... ne méprisez pas ceux qui sont vos maîtres: ils sont égaux à votre mère, en créant les codes de conduite' (premier anak 'ceux qui'; deuxième anak 'ils')

Dans Rāmakerti II (prob. XVIII<sup>e</sup> s.), Sītā exilée dans la forêt appelle l'ascète Vajjamrik qui l'a recueillie: anak tā, lit. seigneur grand-père. Dans les IMA, un des appellatifs des moines est anak yeṇ, lit. mon (ou notre) seigneur.

On notera, d'après ces exemples, que anak a un emploi honorifique, sinon sacré: appellatif et pronom personnel réservé aux princes, aux ascètes et aux moines. On notera, en particulier, l'emploi de l'appellatif anak tā pour les ascètes,<sup>22</sup> qui s'étendra plus tard aux esprits tutélaires locaux. Dans ce dernier cas, l'appellatif se nominalise, d'où khm.mod. anak tā /neak taa/ 'génie tutélaire'.

3.5.3. En khmer moderne, anak a perdu toute valeur nominale. En revanche, sa fonction grammaticale s'est considérablement développée.

3.4.3.1. La fonction d'agent est la plus productive; anak peut être préfixé à différentes sortes de verbes, noms, démonstratifs ou indéfinis. Anak srae (= celui-rizière) est 'le paysan'; anak crīeṇ (= celui-chanter), 'le chanteur'; anak dham (= personne-grand) 'une personnalité';

anak nā (= personne-quelconque) '*celui qui ...*'; anak neh (= personne-ci) '*celui-ci*'.

3.5.3.2. Un doublet graphique de anak, à savoir nāk' (cf. forme nak en vx.khm.), s'emploie aujourd'hui comme classificateur d'êtres humains.

3.5.3.3. Anak est un pronom courtois, 2<sup>e</sup> p., usité dans des énoncés plutôt formels. Anak yal' r̄ de? '*Avez-vous compris?*', est plus courtois que la question familière elliptique yal' r̄ de?

3.5.3.4. Comme appellatif, anak s'emploie pour les femmes, '*Madame, Mademoiselle*', dans le parler standard des communautés urbaines, tandis qu'il s'applique aux hommes d'âge moyen dans certains parlers régionaux.

3.6. Braḥ remonte à vx.khm. vraḥ dont j'ai analysé l'étymologie en détail dans un précédent article (Pou 1976:343-5).

3.6.1. Qu'il nous suffise de rappeler brièvement le mot de base raḥ '*apparaître, briller*', à préfixe labial ayant valeur d'agent, donc '*celui qui est brillant, illustre; l'être sacré*'. Par conséquent, vraḥ désigne en vieux khmer un dieu, le Buddha, un officiant de culte, le roi ou prince, ainsi que la représentation de ces êtres, telle une statue divine.

Vraḥ sert à former des composés à caractère sacré, dont le type est vraḥ pāda '*le roi, Sa Majesté*'. Parallèlement, il peut être préfixé à des noms qui assument, alors, un aspect sacré. J'en relève deux exemples dans une même inscription du X<sup>e</sup> s. (K.956, IC VII, p.130):

vraḥ amā vraḥ pāda kaṃraten ... '*L'oncle du roi ...*' (1.14)

vraḥ kaṃvujadeśa, lit. le saint pays des Kambuja (1.16)

3.6.2. Khm.moy. braḥ a conservé ces valeurs avec de légères modifications.

3.6.2.1. La fonction de préfixe reste inchangée. Par exemple: braḥ dharm désigne la sainte '*Loi*' du Buddha.

3.6.2.2. Dans la désignation des êtres, braḥ tend à se polariser de la façon suivante. D'abord, il s'applique aux êtres sacrés supérieurs qui reçoivent un culte: le Buddha, le dieu Rām, le roi. Dans cet emploi, il peut se rendre par '*le seigneur*', et fonctionne soit comme nom, soit comme pronom. Voici la st.1413 de Rāmakerti I, une phrase dont l'acteur est Khar, démon-roi de Daṇḍakār:

pāñ' braḥ braḥ pāñ' vjñ thkān

cāp' braḥ braḥ crān

buṃ āc nūv dau dīep braḥ

*'Il tire sur le Seigneur, celui-ci tire en retour; il saisit le Seigneur, celui-ci le repousse; il ne réussit pas à s'approcher de lui'.*

3.6.2.3. A l'autre bout de l'axe, braḥ est mentionné dans les IMA comme titre d'un petit dignitaire,<sup>23</sup> parfois préfixé de anak en anak braḥ (cf. aussi Lewitz 1972b:111).

3.6.3. En khmer moderne, 3.6.2.3. est hors d'usage. 3.6.2.1. est inchangé, par exemple: braḥ ādity<sup>24</sup> 'le soleil', braḥ candr 'la lune'.

Quant à 3.6.2.2., braḥ se spécifie comme nom commun, désignant 'le Buddha', à l'exclusion d'autres êtres sacrés. Mais, il ne peut représenter le même être comme pronom; il faut, pour cela, le pronom composé braḥ aṅg (= le saint corps. cf. 2.1.2.2.).

3.7. Lok /lɔk/ est la réduction du sk. loka 'le monde'.

3.7.1. En vieux khmer, loka a la même valeur que jagat, désignant soit 'le monde' dans son ensemble, soit 'les personnes'.

3.7.2. En khmer moyen, lok, dans le sens partitif, se spécialise dans sa fonction honorifique. Il ne s'agit pas seulement de 'gens', mais de 'gens de bien', des personnes qu'on doit respecter, vénérer, dont voici quelques illustrations:

doḥ nūv oy rājy bhirut grañ

sot sam mān hmañ

ṭaṃnTel nai lok nindā (Rāmakerti I, st.282)

*'Si nous confions le royaume à Bhirut, (dit Dasarath), nous commettrons une injustice, et nous serons blâmés par les gens de bien'.*

lok thā bhoen bhlī

maen bit moḥ fī

buṃ smoe suriyā ... (Kūn cau, st.9)

*'Les sages disent que le feu, bien que réellement lumineux, ne peut égaler le soleil ...'.*

Grâce à ces deux exemples, on saisit la valeur générique de lok en khmer moyen: ce sont 'des gens de bien' ou d'élite, remarquables par leur savoir, leur talent, leur vertu ou leur sagesse. Un autre exemple, tiré le Lpoek Aṅgar Vatt, va confirmer cette définition du mot lok:



lok chlāk' ʔoy nūv jamboḥ

ʔoy kĭen ʔoy koḥ

ʔoy jruñ ʔoy jroy phañ nā (st.201)

*'Avec conviction, on a sculpté (des scènes) à tous les coins et à tous les angles (des panneaux de murs)'*.

J'ai traduit par 'on' le mot lok du début de la strophe qui, indubitablement, se rapporte aux artistes du passé qui ont droit à la vénération de la postérité.

3.7.3. Par conséquent, lorsque lok devient appellatif en khmer moderne; il s'applique à certaines entités sacrées, comme le ciel et les astres, et aux gens 'supérieurs' par leur âge, leur condition sociale ou intellectuelle. C'est par ce biais qu'il est usité comme pronom personnel 2<sup>e</sup> et 3<sup>e</sup> p. honorifique.

3.8. Me et pā, bien que d'aspect vétuste dans la cadre de notre étude, mérite un nouvel examen.

3.8.1. Deux noms à l'origine, ils signifiaient respectivement 'la mère' et 'le père'. Me se rencontrait sous les formes de me ou ame, tandis que pā était au départ un dissyllabe noté vappā (cf. *infra*). Les deux mots accouplés en un composé désignaient naturellement 'les parents':

bhūmi me vapā añ ti uttara ti jvan ta rūpa ame añ (K.693, IC V,B, 1.19-20)

*'la terre de mes parents, au Nord, fut offerte à l'image de ma mère'.*

D'autre part, me était appellatif de femmes d'un certain âge, tandis que vāp, forme tronquée de vappā, était un appellatif courtois d'hommes (cf. Lewitz 1976).

3.8.2. Après la grande lacune des XIV<sup>e</sup> et XV<sup>e</sup> siècles, on voit apparaître de nouveau les deux mots, modifiés.

3.8.2.1. Leurs formes sont réduites exclusivement à me et pā (stabilisées jusqu'à l'heure actuelle). Comme en vieux khmer, ce sont encore deux lexèmes signifiant séparément 'la mère' et 'le père', ou collectivement 'les parents'. La st.8 du Kram dit:

āsūr me pā

ciñcim raksā

buṃ oy antarāy ...

*'Pensez à vos parents qui vous ont élevés et protégés ...'*

3.8.2.2. Mais c'est la fonction d'appellatifs qui se développe notamment en khmer moyen, et qui fixe l'emploi de ces termes comme pronoms personnels. Pā précédé de jī (cf. 2.2.3.1.), donc jī pā, s'applique au 'père' comme appellatif et pronom 2<sup>e</sup> p., alors que, employé seul, il s'applique aux jeunes gens. Me est un appellatif courant de femmes (cf. IMA), et s'emploie comme pronom 2<sup>e</sup> p. A titre d'exemple, citons Rāmakerti I, où Rām utilise me à l'adresse de Sītā, et pā à celle de Laks.

3.8.3. Le khmer moderne a conservé les usages du khmer moyen, en en modifiant les nuances.

3.8.3.1. Les noms me /mɛɛ/ et pā /baa/, pour les êtres humains, est d'un usage rare et chargé de valeur morale, sinon mystique. En revanche, ils sont d'un usage courant lorsqu'il est question d'animaux. Me cĕm (= mère mouton) est '*la brebis*'; pā go ou go pā est '*le taureau*'.

3.8.3.2. L'évolution est très marquée dans les fonctions d'appellatifs. Pā est presque hors d'usage, excepté dans quelques coins de la campagne. Quant à me, appliqué aux êtres humains féminins, il prend une valeur dépréciative telle qu'il devient un préfixe injurieux.

3.8.3.3. Par ailleurs, me seul prend le sens de '*maître, chef*'. Il implique l'idée de 'supériorité, suprématie et grandeur'. Me phdah (= me + maison) est '*la maîtresse de maison*'; me dāhān (= me + soldats) est un '*officier de l'armée*'; me ɬai (= me + main) est '*le pouce*'. Cet usage s'étend aux verbes, pour exprimer un superlatif dans une formation ouverte: me khjil '*très paresseux*', me soec '*aimer beaucoup rire*'.<sup>25</sup>

Cette analyse détaillée des termes servant de pronoms personnels en khmer moderne ne fait que confirmer leur ambivalence catégorielle et fonctionnelle. Or, cette ambivalence est réelle, effective, puisqu'elle est le résultat d'une longue évolution historique d'un groupe de lexèmes dont le concept de base était lié à 'la personne', avec ses fonctions et les qualités particulières que lui attribuait la société. Evolution, néanmoins structurée, qui a abouti à un système dont on aura un aperçu dans l'Appendice suivant.

N O T E S

1. Telle qu'elle était usitée au Cambodge jusqu'en 1975.
2. Ce problème, loin d'être particulier au khmer, se retrouve dans de nombreuses langues asiatiques qui me soient connues.
3. Graphie commune aux signifiants /kərunaa/ et /kənaa/, à savoir karuṇā: d'où, khñuṃ braḥ karuṇā et khñuṃ karuṇā.
4. Ce caractère aberrant est encore plus accentué par l'ellipse des termes de grammaire dans le parler familier - autre trait commun à nos langues - laquelle dépend étroitement de la valeur informative du contexte. Un locuteur peut donc choisir d'omettre les termes de grammaire dans ses énoncés, et le plus fréquemment les pronoms personnels.
5. A comparer ces deux énoncés:
  - a) pañ jhṭ 'mon aîné(e) est malade'. Pañ est indubitablement un nom-sujet.
  - b) pān dhvœ dau 'Faites-le!', énoncé injonctif sans équivoque, marqué par la particule post-verbale injonctive dau. Pañ y est un appellatif employé comme pronom, 2è p.
6. J'élimine de la liste de H.J. Pinnow le pronom cité sṛ̃ṇ [seng]. En effet, déictique anaphorique en vx.khm. et en khm.moy., il est devenu un adverbe en khm.mod., 'presque'.
7. Je rappelle que ce qu'on nomme 'adjectifs possessifs' dans les manuels sont en réalité des pronoms personnels placés après les mots qu'ils déterminent.

8. Khñuṃ pād /knom baat/, dans un énoncé courant, rapide, devient simplement /mbaat/.

9. En khmer moyen, hañ s'appliquait aux hommes comme aux femmes.

10. Cet exemple d'expansion lexicale illustre l'élaboration du rā-jasabd, ou '*vocabulaire royal*', en système, dont j'entreprendrai l'étude prochainement.

11. Terme employé également comme classificateur d'êtres sacrés. A noter que le khmer moyen de Rāmakerti I révèle un terme semblable, braḥ tanū (= le saint corps) '*il, lui, soi*' (cf. sk. tanū '*le corps*'), qui est tombé en désuétude.

12. Vx.mōn ey /ʔɔy/, mōn mod. ay /ʔoa/; pal. ɔ; rlang, o; danaw, o; bah., eŋ, iŋ; srê, aŋ; maa, aŋ; chrau, aŋ, iŋ; hal., aw; jeh, aw; sed., ə, au; khmu, oʔ; chong, eŋ; lawa, au; asl., iŋ, ic, əŋ; munda, aŋ, iŋ, iŋ.

13. cf. la st.7 de Cpāp' prus (prob.XVIII<sup>e</sup> s.), Jenner and Pou 1976: 332.

14. Pal., ye; wa, ɛ; khas., ŋii; asl., heʔ, yeʔ; pacoh, he; mnong g., hii; sed., ŋin, ɲeɛn; bah., noon; jeh, hal., yuan, yon.

15. Un doublet de vā /wia/ et vā /waa/, préfixé souvent de ā- en /ʔa waa/, appellatif de petits garçons (régional).

16. Parfois, le pronom añ, dans la bouche de Rām, est adouci par l'adjonction d'un terme de tendresse, par exemple añ pañ '*moi (qui suis) ton aîné*'.

17. cf. khm.mod. khnuṃ mcās', 2.1.2.1.

18. Dans certaines campagnes, les hommes utilisent volontiers le pronom ge '*vous, tu*' en s'adressant à leurs épouses. Mais ici, le VK impose l'orthographe geh (1968:140a), car, dit-il, ce pronom provient "du sk. cu p. geha qui signifie '*la maison, la maîtresse de maison (sic)*'". Il est plus judicieux de renoncer à cette étymologie forcée pour nous en tenir à vx.khm. ge '*personne*' > '*vous, tu*'.

19. Comme pour perpétuer la tendance ambivalente, la langue familière moderne admet ā pour les filles, et nān pour les jeunes garçons!

20. Dans mon article (Lewitz 1976:761-71), je n'ai pas noté un homologue moderne de vx.khm. si. Je vais donc combler cette lacune avec ptī sī /pdey sey/ 'le mari', où sī n'est pas un simple élément assonant, mais le mot khm. pour 'homme, mari', couplé à son synonyme ptī provenant de sk. pati 'mari'.

21. Le khm.mod. exige ge à la place de anak dans ce vers.

22. Dans cet emploi, anak tā est maintenant remplacé par lok tā. cf. la valeur de lok plus loin: 3.7.

23. Dépréciation de nombreux titres à l'époque moyenne.

24. Ādity /ʔatit/, non préfixé de brah, signifie 'la semaine'. Les savants de l'IB, jugeant cet emploi 'impropre', ont décidé de le rayer en créant saptāha: /sapətaahaʔ/ qui n'a jamais pris racine en khmer.

25. Dans ces cas de composés, il faut noter comme substitut de me- le sk. mahā- (dès le khmer moyen). Ainsi, on peut dire mahā khjil, mahā soec, etc.

## APPENDICE

## Tableau des pronoms personnels usuels du khmer moderne

## A. 1ère PERSONNE

khñuṃ /knom/: pr. courtois, sg. ou pl.

khñuṃ pād /knom baat/: pr. très poli, sg. ou pl., employé par les hommes.

yoeñ /yəəŋ/: pr. pl. courant; pr. sg. dans la langue familière.

yoeñ khñuṃ /yəəŋ knom/: pr. pl. poli.

añ /aŋ/: pr. sg., employé vis-à-vis de soi-même ou des inférieurs.

añ khñuṃ /aŋ knom/: pr. sg. poli.

khñuṃ mcās' /knom mcah/: pr. sg. ou pl., employé par les femmes s'adressant à des princes.

khñuṃ braḥ karuṇā /knom preah kərunaa/: pr. sg. ou pl., employé par les hommes s'adressant au roi; /knom kənaa/, pr. sg. ou pl., masc. ou fém., employé par les laïcs s'adressant à un moine.

ḍūl braḥ paṅgaṃ /tuul preah bəŋkum/ ou /tuul bəŋkum/: pr. sg., employé par les hommes s'adressant à des princes.

ātma /ʔatmaa/: pr. sg. employé par les moines.

ḡnā /kniə/: pr. sg. de la langue familière.

ḡe /kɛɛ/: pr. sg., moins poli que le précédent.

əñ /ʔaəŋ/: pr. sg., encore moins poli que le précédent.

## B. 2è PERSONNE

əñ /ʔaəŋ/: pr. sg. ou pl. peu poli.

ne əñ /mɛɛ ʔaəŋ/: pr. sg. ou pl. dépréciatif pour les femmes.

hañ əñ /həŋ ʔaəŋ/ ou /ʔəŋ ʔaəŋ/: pire que le précédent, presque injurieux.

ā əñ /ʔaa ʔaəŋ/: pr. sg. ou pl., de même type que les deux précédents, mais pour les hommes.

- pā /baa/: pr. sg. ou pl. de la langue littéraire, et de certains parlars régionaux, destiné aux jeunes garçons (rare).
- cau /caw/: même type que le précédent, mais appliqué aux jeunes gens des deux sexes.
- anak /neak/: pr. sg. ou pl. courtois et neutre dans la langue courante; pr. familier pour les hommes, dans certaines campagnes; pr. très poli pour les femmes.
- nān /niən/: pr. sg. ou pl. poli, employé pour les femmes plus jeunes que soi, ou pour les enfants des deux sexes.
- lok /lɔk/: pr. sg. ou pl. honorifique pour les hommes.
- brah aṅg /preah ʔŋ/: pr. sg. pour le Buddha, une divinité, ou le roi.
- brah tej brah guṇ /preah dac preah kun/: pr. sg. ou pl., employé pour les dignitaires.
- gāt' /kɔət/: pr. sg. familier.
- khluon /kluən/: pr. sg. peu poli, cavalier.
- ñom /nɔm/: pr. sg. ou pl. appliqué par les moines à leurs interlocuteurs laïques.

C. 3<sup>e</sup> PERSONNE

- gāt' /kɔət/: pr. sg., rarement pl., courant.
- ge /kɛɛ/: pr. pl. courant.
- vā /wiə/: pr. sg. ou pl., employé pour les enfants, les animaux; déictique de la langue familière.
- brah /preah/ ou brah aṅg /preah ʔŋ/: pr. sg. ou pl. pour les êtres sacrés.
- brah karuṇā /preah kərunaa/: pr. sg. pour le roi.
- stec /sdac/: pr. sg. pour le roi.
- samtec /səmdac/: pr. sg. pour la reine ou un grand prince.
- lok /lɔk/: pr. sg. ou pl. pour les personnes d'un haut rang.
- gnā /kniə/: pr. sg. ou pl. peu courtois.

TEXTES CONSULTÉS

AYMONIER, Etienne François

1878        *Textes khmers*, 266-96. Saigon:

COEDÈS, George, ed.

1937-66    *Inscriptions du Cambodge [IC]*. 8 vols. Hanoi-Paris:  
Imprimerie d'Extrême-Orient.

JENNER, Philip N. and Saveros POU

1976        'Les Cpāp' ou 'codes de conduite' khmers: II. Cpāp'  
prus'. BEFE0 63:313-50.

LEWITZ, Saveros

1968        'L'accentuation syllabique en cambodgien'. In: *Papers*  
*of the CIC Far Eastern Language Institute*, 155-67. Ann  
Arbor.

1970        'Textes en Kmer moyen: Inscriptions modernes d'Angkor 2  
et 3'. BEFE0 57:99-126.

1971        'Inscriptions modernes d'Angkor 4, 5, 6 et 7'. BEFE0  
58:105-23.

1972a        'Inscriptions modernes d'Angkor 1, 8 et 9'. BEFE0 59:  
101-21.

1972b        'Les inscriptions modernes d'Angkor Vat'. *Journal*  
*Asiatique* 260:107-29.



- 1973a 'Inscriptions modernes d'Angkor 17, 18, 19, 20, 21, 22, 23, 24 et 25'. BEFE0 60:163-203.
- 1973b 'Inscriptions modernes d'Angkor 26, 27, 28, 29, 30, 31, 32, 33'. BEFE0 60:205-42.
- 1974 'Recierches sur le vocabulaire cambodgien (VIII): du vieux Khmer au Khmer moderne'. *Journal Asiatique* 262: 143-70.
- 1976 'Note on Words for Male and Female in Old Khmer and Modern Khmer'. *Austroasiatic Studies* 761-71. Honolulu.

## PINNOW, H.J.

- 1965 'Personal Pronouns in the Austroasiatic Languages: A Historical Study'. (Translated by H.L. Shorto.) *Lingua* 14:3-42.

## POU, Saveros

- 1974 'Inscriptions modernes d'Angkor 35, 36, 37 et 39'. BEFE0 61:301-37.
- 1976 'Recherches sur le vocabulaire cambodgien (IX)'. *Journal Asiatique* 264:333-55.
- 1977a 'Inscriptions en khmer moyen de Vat Athvéa (K.261)'. BEFE0 64:151-66.
- 1977b *Rāmakerti (XVI<sup>e</sup>-XVII<sup>e</sup> siècles)*, traduit et commenté par ... Publications de l'EFEO, vol.CX.
- 1979 *Rāmakerti (XVI<sup>e</sup>-XVII<sup>e</sup> siècles)*. Texte khmer publié par ... Publications de l'EFEO, vol.CXVIII.
- to appear *Rāmakerti II*, presented and commented upon by S. Pou. In preparation.

## POU, Saveros and Philip N. JENNER

- 1975 'Les Cpāp' ou 'Codes de conduite' khmers, (I) Cpāp' Kerti kāl'. BEFE0 62:369-94.

- 1977 'Les Cpāp' ou 'Codes de conduite' khmers, (III) Cpāp' Kūn  
cau'. BEFE0 64:167-215.
- 1979 'Les Cpāp' ou 'Codes de conduite' khmers, (V) Cpāp' Kram'.  
BEFE0 66:129-60.

## STANDARDISATION AND PURIFICATION: A LOOK AT LANGUAGE PLANNING IN VIETNAM

NGUYỄN ĐÌNH-HÒA

0. Language planning has been defined as "the organized pursuit of solutions to language problems, typically at the national level" (Fishman 1975:55). In the second part of his book on *Language and Nationalism: Two Integrative Essays*, Fishman (1975:40-85) discusses the impact of nationalism on language and language planning by analysing the role of language as "the link with the glorious past", as "the link with authenticity" and as "part of the message of nationalism" (pp.44-5). The Vietnamese people have always been conscious of the real status and function of their language as symbol of the unity of the nation. Under colonial rule, although using French as the medium of instruction, school-children were encouraged to speak unadulterated Vietnamese and avoid the macaronic language consisting of a mixture of Vietnamese and French (Nguyễn Đình-Hoà 1975:44-5). French-trained teachers promoted the teaching of the Vietnamese script to illiterate adults through the Association for the Dissemination of Quốc-ngữ and the publication of numerous books and periodicals where nearly every technical subject could be discussed in the national language clothed in the Roman script (Nguyễn Khắc-Kham 1976:189-90).

Even prior to the August 1945 revolution, the Indochinese Communist Party in its Cultural Program (Đề-cương văn-hoá Việt-nam) issued in 1943 raised the question of making the task of "struggling about language" an urgent one, "which should aim at unifying and enriching the national language, determining grammatical rules and improving the quốc-ngữ script" (Vương Lộc 1975:21). Then after the August revolution, in the midst of this anti-French hostilities, Trường Chinh called upon everybody to "create a movement to Vietnamize both language and literature" and "to resolutely defend our mother tongue" (T. Tr. 1948).

During the partition of the country, linguistic scholars in both halves recognised that "to preserve the purity and clarity of the Vietnamese language and to make it richer, more beautiful and more precise every day is the task of every citizen of Vietnam, and especially of linguists, writers and newspapermen." (Lưu Văn Lăng et al 1961:83).

This paper examines steps in the processes of language planning in the reunified Vietnam.

1. Language planning in Vietnam seems to have moved beyond the stage of "policy formation" (Jernudd 1973:15) or "norm selection" and well into the stages of "codification and elaboration" and "implementation"<sup>1</sup> (Haugen 1966; Fishman 1972:56-7).

1.1. It can be said that language policies were formulated officially only in 1966 when three meetings were organised between the end of January and the beginning of February by the Writers Association (Hội nhà văn Việt-nam) and the Institute of Literature (Viện văn-học) to discuss the problem of preserving the purity and clarity of the Vietnamese language. As reported in the organ of the latter institute, participants included poets, writers, newsmen, research workers in literature, linguistics and other sciences in the Institute of Social Sciences, as well as college teachers, representatives of the press, news agencies, radio stations and publishing houses. At these meetings led by Social Sciences Committee Chairman Nguyễn Khánh Toàn, Minister of Education Nguyễn Văn Huyền and Vice Minister of Education Lê Liêm, there was a lively exchange of ideas and opinions. Premier Phạm Văn Đồng attended all three meetings, and his remarks given at the last meeting were later published in *Tạp-chí Văn-học* 75 (March 1966:1-5, 93-5), then reproduced in the theoretical journal of the Lao-động (Workers) Party *Học-tập* 123 (April 1966:7-13).

The follow-up was reported in *Tạp-chí Văn-học* 80 (August 1966:106): nearly 60 linguistic workers held their own meeting on June 3, 1966 to discuss the preservation of the purity and clarity of Vietnamese. These scholars from the Institute of Social Sciences, the University of Hanoi, Hanoi Teachers College, Vinh Teachers College met under the chairmanship of Professor Đặng Thai Mai, President of the Institute of Literature, Professor Hoàng Xuân Nhị, Dean of Letters at the University of Hanoi, with other staff members of the Educational Sciences Institute, the School of Foreign Languages, etc.

1.2. More recently books and articles have been devoted to the same problem, which was set forth officially for the first time in 1966 with the participation of the Premier himself.

Speaking as a writer of great sensitivity, Premier Phạm Văn Đồng pointed out that the Vietnamese language being rich and beautiful, its speakers must be "deeply conscious of its richness and beauty and of its unlimited capabilities" (Phạm Văn Đồng 1966:3). He said (1966:4) that "language must reflect life and serve life, the daily life and the revolutionary struggle of the masses, our people's present anti-U.S. struggle for national salvation, political, cultural, literary, and artistic life, and scientific and technological life." The Premier first reviewed how since 1945 Vietnamese has been used in every field - political, economic, social, cultural, artistic, scientific and technical - mentioning the liquidation of illiteracy within a short time and the widespread use of Vietnamese as the medium of instruction, even at the college level, for science and technology. The Vietnamese leader said that Vietnamese had been made into an excellent tool in the cultural and ideological revolution. Then he cited three areas where Vietnamese must concentrate their initiative, display sensitivity and perseverance and work firmly with planning and organisation to sustain efforts at purism: (1) to preserve and develop the vocabulary; (2) to respect the grammar of (spoken and written) Vietnamese; and (3) to maintain the originality and essence of the national language in every genre of writing.

As far as vocabulary is concerned, he warned against a "communicable disease" which the late President Hồ Chí Minh used to criticise - the tendency or phenomenon that consists in using too many foreign loanwords, especially from Chinese:

xạ kích	instead of	bắn súng	'to shoot, fire'
cự-ly		khoảng cách	'distance'
tiến-hành		làm	'to do'
sử-dụng		dùng	'to use'
phát-biểu		nói	'to speak'

In his important speech, the Premier touched on the elaboration of scientific terminology. Technical terms have to be introduced systematically, he said, in all disciplines. However, who is going to be responsible for terminology work and how the new jargon can be worked out are important issues according to him.

It is significant that the Premier used the native words vốn chữ instead of từ-vựng for 'vocabulary, lexicon' and phép tắc instead of ngữ-pháp for 'grammar'.

Grammar is a very important subject, the Vietnamese leader explained, "because we teachers need it, our students need it, and foreigners who want to learn Vietnamese also need it." (1966:93). Here his major contribution was a warning against using the grammar of a foreign tongue to write a grammar of Vietnamese. He advised grammarians to start from "the realities of Vietnamese itself" just as "the natural conditions of our country and the revolutionary and social realities of our people" should help solve "such problems in our present life as growing rice, building a home, making a dress, cooking a meal, etc." (1966:93)

Thirdly, in making Vietnamese even richer and more beautiful and helping the language "keep up with the revolution and our people's rich life", people should try to use new expressions, new phraseology - for instance, nominalisation - yet at the same time ensuring that the language will be developed firmly on the foundation of the old lexicon and retain its style, its originality, its quintessence."

Language planning usually involves a "body of experts specifically delegated the task of preparing a plan" and who "ideally estimate existing resources and forecast." (Jernudd and Das Gupta 1971:196)

In his conclusion Premier Phạm Văn Đồng turned to his "body of experts" in the audience; he urged a division of labour, first with the Linguistics Section of the Institute of Literature - it has since grown into an Institute of Linguistics independent of, but working closely with, the latter institute - compiling a general dictionary and different glossaries of technical terms, in consultation with specialists in various branches of learning, then secondly the linguists also writing a grammar of Vietnamese, with the active participation of writers, and finally intensified efforts on the part of writers and journalists through the Writers' Association (Phạm Văn Đồng 1966:94).

1.3. The Premier noted the presence of educational leaders Nguyễn Văn Huyền and Lê Liêm at the 1966 sessions: since the school shoulders a heavy responsibility, he asserted that "the young generation must speak and write better than we do." (p.95)

The official language policy was thus clearly enunciated: "This is a long-range undertaking. We must be patient, look far ahead and have a broad view of the problem, proceed step by step with all our sense of responsibility, with a pride in our national language, with enthusiasm and confidence in our contribution to a task that is both important and extremely beautiful." (Phạm Văn Đồng 1966:95)

Dr Nguyễn Văn Huyền, then Minister of Education, analysed the weaknesses of the language arts program. First graders had not been taught

(1966) "carefully, minutely and fully how to read, write and speak" their mother tongue, and in Level II and Level III schools, more emphasis had been given to Vietnamese literature than to Vietnamese language, with the textbooks not serving (p.7) as "models in every respect".

Effective use of the language, Minister Nguyễn Văn Huyền said, cannot be attained (1966:8) "through prohibitions and orders, but through persistent training and time-consuming education." He also stressed prevention instead of cure, and he explained that children learn not only in their classrooms, but also from "the speech of the people, from proverbs and sayings, folk songs, pieces of classical prose and poetry, and model works by our writers and poets." (Nguyễn Văn Huyền 1966:95)

1.4. It is significant that in the same issue of the literary review, there are articles on the same topic written by such literary figures as Xuân Diệu (9-20, 27), Nguyễn Tuân (21-7), and Chế Lan Viên (28-32). The use of "good" Vietnamese in the press (Quang Đạm 1966:60-71, 91-8), on the stage (Nguyễn Bổng Hồ 72-5) and in movies (Vũ Bá Hùng 75-80) was also discussed at length in that issue.

2. The prescriptivist or normative tendencies in the codification stage are evident in book titles - *Rèn luyện về ngôn-ngữ* 'Linguistic training' (Nguyễn Kim Thản 1966), *Nói và viết đúng tiếng Việt* 'Speak and write Vietnamese correctly' (Nguyễn Kim Thản et al 1967) - and in writings on the specific topic of standardisation (for instance, Nguyễn Hàm Dương 1974 and 1975).

A standard language has been defined as "a codified form of language, accepted by and serving as a model to, a larger speech community." (Garvin 1973:25). To the same definition Garvin added that "a given language situation can be described as meeting the criteria for a standard language to a given degree, rather than absolutely." (Garvin 1973:25). Elsewhere Garvin and Mathiot (1956:788) point out that the functions of a standard language (unifying and separatist functions, prestige function and frame-of-reference function) occasion a set of cultural attitudes such as language loyalty, pride and awareness of the norm.

In the case of Vietnam the search for what constitutes the standard to follow has been carried out seriously and diligently revolving around the problem of evaluating competing forms. The first volume of a work entitled 'Studies in Linguistics' (*Nghiên-cứu ngôn-ngữ-học*) and published in 1968 was devoted to this subject. Issued under the auspices of the Linguistics Section of the Committee on Social Sciences, it contains articles dealing with the theme of linguistic purity and clarity, as

shown in the subtitle, thus touching on standardisation, spelling reforms, abbreviations, semantic change, abuse of Chinese loanwords and dialect lexicon.

As the debate was launched in earnest in 1966, one writer said that it is not easy "to distinguish what is correct and what is incorrect, what makes Vietnamese richer and more beautiful and what 'corrupts' it." (Quang Đạm 1966:60). The poet Chế Lan Viên points out that there are indeed "many different concepts of purity and clarity, hence many different evaluations." (Chế Lan Viên 1966:29). A famous essayist aware of the consequences of *laissez-faire* asks the question how to conceive the problem so as not to prevent "each writer from having his own style, his own features?" (Nguyễn Tuấn 1966:26). Another relevant question has also been asked: How to avoid "the situation wherein one relies only on one's impressions?" (Hoàng Tuệ 1961:55).

In answer to all these questions, Hoàng Phê, presently Associate Editor of the review *Ngôn-ngữ* 'Language' - organ of the Institute of Linguistics - quotes Hall (1950), Bloomfield (1933) and Martinet (1961) and says "it is wrong to oppose subjectivism and fall into naturalism." (Hoàng Phê 1968:7-8). He claims that "linguistic phenomena are evaluated" and that "their evaluation leads to a choice of expression, to creativity in speech and to different attitudes toward new phenomena in language" (Hoàng Phê 1968:8). Claiming that you cannot "leave your language alone" and that the crux of the matter is not whether to evaluate, but how to evaluate (1968:9), Hoàng Phê underlines that "standardisation is the central concept in the problem of preserving the purity and clarity of language" (1968:16), and he points out the obligatory (p.17) and stable (p.22) character of standards as well as their relativity (Hoàng Phê 1968:26).

"New standards are taking shape and have to be built up. Educated Vietnamese is moving toward a new purity and clarity, a higher purity and clarity, a purity and a clarity that stress not only form but also content, that not only follow the old crease, but is also creative, that is not monotonous, but rich, not affected, but simple, not rigid, but developing." (Hoàng Phê 1968:34). Likewise, both Lê Xuân Thai and Hồng Dân discuss standard and development in terms of individual creativity (Lê Xuân Thai 1968:42-52) and of "the sum of rules governing the use of a language in a given period" as constituting the standard of that language (Hồng Dân 1968:53).

Standardisation in Vietnam thus consists basically of two steps: first a model for imitation is created, then that model is promoted over rival models (Ray 1963:70). In two articles devoted to standardisation



specifically, Nguyễn Hàm Dương notes the importance of vigilance and caution needed to forge the model language as the tool of communication and social action:

"To preserve the purity and clarity of Vietnamese is not only to polish one's style, to weigh each word or to be careful in speaking and writing. These are necessary, but not sufficient conditions. To preserve the purity and clarity of Vietnamese is to struggle for the concision, sharpness, richness, exemplariness and the strong expressiveness of Vietnamese, raise the effectiveness of its use in every field of endeavor in society, fully expand its potentialities, make it really an important weapon of the socialist culture, and a means of communication and thinking of tens of millions of Vietnamese." (Nguyễn Hàm Dương 1975:26-7).

It is interesting to note that the same linguist, in an earlier article (Nguyễn Hàm Dương 1974), uses *tiêu-chuẩn-hoá* for 'standardisation' but later switches to the shorter form *chuẩn-hoá* (Nguyễn Hàm Dương 1975). The very term for 'linguistics' preferred in pre-1975 technical discussions in South Vietnam was *ngữ-học*, but now the longer form *ngôn-ngữ-học* has gained currency.

After saying that "the content of the problem of preserving the purity and clarity of Vietnamese is closely linked to the task of standardisation in this present revolutionary period", the two concepts being interrelated (1975:27), Nguyễn Hàm Dương enumerated the practical tasks as follows: (1) to conduct a linguistic survey covering both the spoken and the written languages (p.31); (2) to establish a language policy built on the interaction between social elements and the internal structure of language, since such a policy is essential to the selection and evaluation processes (p.32); (3) to study the language of the late President Hồ Chí Minh (p.32); and (4) to study model writings by representative writers and poets (p.32).

Implementation tasks do not neglect such effective means as language teaching in the schools, where there is the conscious intervention of society in the development of language, and the mass media (press, radio, television, theatre, movies, etc.) (p.33).

The production of such tools as dictionaries and standard grammars (Nguyễn Hàm Dương 1975:34) is the responsibility of linguistic scholars, since "now more than ever the central task of linguistics is still the preservation of the purity and clarity of the Vietnamese language" as Social Sciences Committee Chairman Nguyễn Khánh Toàn urged in a message on the "great turning point and linguistics (1973:1-4).

As implementation shifted into high gear, meetings and seminars were organised. A full conference on standardisation was convened by the Linguistics Department of the University of Hanoi on February 27 and 28, 1975. On April 4, 1975 the Institute of Linguistics (Viện Ngôn-ngữ-học) itself held a meeting to hear eleven reports on the efforts.

3. Since the planners are keenly aware of the respective roles of speech and writing, the function of the latter as the medium of communication "between speakers separated in time and space" and "its permanence and importance for the community" "permit and require a different kind of treatment from that which is accorded natural speech." (Haugen 1966: 53). This explains why script improvement and spelling reforms have received a great deal of attention in the process of language planning.

The Cultural Program issued by the Party in 1943 already mentioned the need for spelling reforms based on recommendations found in publications (*Khoa-học, Tri-tân, Thanh-nghị*, etc.) launched by French-trained Vietnamese professors.

The Second Cultural Congress convened in 1943 further pursued the idea of orthographic improvement. Then, six years after the partition was decided at the Geneva Conference, 67 participants held discussions on the improvement of quốc-ngữ. The proceedings of this conference were published the following year (*Hội-nghị cải-tiến quốc-ngữ* 1961), and the underlying principles of reforms proposed in the major conference papers were as follows:

- (A) Spelling reforms must be "based on the presently used Roman script" and "implemented step by step."
- (B) The reforms must be bold, but cautious and firm at the same time;
- (C) Each step should prepare for the next step;
- (D) The reforms must aim at making the quốc-ngữ script more rational and at the same time more simple and more adequate. (p.113)

One paper stressed that the Conference "must confirm the standard pronunciation of Vietnamese before studying spelling reforms and orthographic uniformization." (p.120). Another participant added that efforts should be made to "avoid unnecessary upheavals." (Bùi Công Trùng 1961: 383).

Professor Nguyễn Kim Thản of the Institute of Linguistics summed up by pointing out (1961:212ff) that the Conference should make "the writing system follow the development of the Vietnamese language more

closely, and at the same time pay attention to the characteristics of our national language." Thus, in the draft program inserted as an appendix of the Proceedings volume (*Hội-nghị cải-tiến quốc-ngữ* 1961: 1141-95), it is clear that although the reforms recommended were based on the Hanoi dialect, regional differences reflected in variant spellings *ch* : *tr*, *x* : *s*, *d* : *gi* : *r* were respected. There were also recommendations regarding the dropping of *h* in the transcription of initial velars, presently *gh-* and *ngh-*, before front vowels. Consistent use of initial *k-* was also recommended. The letters *f-* and *d-* should replace the present graphemes *ph-* and *đ-* (the so-called "barred *d*") respectively. Alien clusters such as *br-*, *cl-*, etc. would be used in technical and scientific terms.

The Conference designated a Committee of Seven to recommend spelling reforms. Professors Đặng Thai Mai and Phạm Huy-Thông served as Chairman and Vice Chairman, respectively. The other members were Nguyễn Như Kontum, Nguyễn Quang Hưởng, Nguyễn Kim Thản, Hoàng Tuệ and Hoàng Phê (Secretary of the Committee). In February 1961, upon his return from a trip abroad Professor Nguyễn Tài Căn was also invited to join the group, but in a letter dated October 20, 1960, on behalf of the Committee, Chairman Đặng Thai Mai invited teachers and linguists in South Vietnam to lend their co-operation (*Hội-nghị cải-tiến quốc-ngữ* 1961:389-92).

A second Conference on the *quốc-ngữ* script was convened in 1962 with 51 participants discussing the draft program recommended by the original Committee of Seven. On May 5, 1962, in a second open letter to his colleagues in the south Đặng Thai Mai wrote that he hoped "northern and southern intellectuals will be able to meet in a national conference to discuss the problem of language and writing as well as all other cultural problems of our country."

Let us note that in South Vietnam, the Cultural Institute (Viện Văn-hoá), established on July 8, 1966, had in addition to other functions the charge of compiling dictionaries of the Vietnamese language, encyclopedic dictionaries, and of preparing a grammar of Vietnamese (Decree No.125-LS/VHXX). A Committee on the Codification of Language and Writing (Ủy-ban diễn-chế văn-tự) set up earlier had the main thrust of promoting the production of technical glossaries in the natural and physical sciences. Hand in hand with "considerable instruction in linguistics" at institutions of higher learning in South Vietnam, research projects were carried out by a number of specialists, whose activities are mentioned in a rather complete survey (Thompson and Thomas 1967:815-46) of the field of Vietnamese linguistics as of 1964.

Judging from technical discussions often reproduced or at least summarised in journal articles, the spelling reforms would (a) improve the transcription of some vowel, consonant and semivowel phonemes; (b) create new syllables needed in both transliteration and translation; (c) run together two or more syllables in compound words and thus eliminate the often erratic use of hyphens.

The efforts of the widely respected scholar Lê Ngọc Trụ in compiling an orthographic dictionary, *Việt-ngữ chính-tả tự-điển*, which was re-printed in 1973, were emulated in North Vietnam, where teamwork seems to be the rule and not the exception in and around the Institute of Linguistics, which held a conference, June 9-10, 1971, just to discuss punctuations and capitalisation. The draft proposal on capitalisation (*Dự-thảo qui-tắc viết hoa*) and two articles discussing it (Dương Lan Hải 1972:53-60 and Phan Thiệu 1972:61-6) were published in No.11 (March 1972) of the journal *Ngôn-ngữ*, pages 49-66, to help dissemination.

4. Preoccupation with the treatment of foreign loanwords has been constant. Vietnamese has borrowed less from French and English than from Chinese (Nguyễn Đình-Hoà 1975:36-59, Nguyễn Đăng Liêm 1975:15-35).

The late multilingual President Hồ Chí Minh used to admonish government employees and cadres for excessively borrowing from the Chinese language. He warned writers and reporters alike against borrowing to the point of not being understood by the masses; he prescribed the use of native Vietnamese words instead of Sino-Vietnamese terms, thus

ba tháng	instead of	tam-cá-nguyệt	'quarter(ly)'
xem xét		quan-sát	'to observe'
đánh vào sâu		tung-thâm	'strike in depth'
xe lửa		hoả-xa	'train'
máy bay		phi-cơ	'aeroplane'
giúp đỡ		hỗ-trợ	'to assist, help'

Such terms as vùng trời 'air space', preferred to không-phận, chữ thập đỏ 'Red Cross', preferred to Hồng-thập-tự (Nguyễn Đình Hoà 1975:50) were first suggested by the late President, who called immoderate borrowing "a disease" and even made fun of the new literates who incorrectly used Chinese-borrowed lexemes, being ignorant of their etymology.

While encouraging the use of completely Vietnamised elements, President Hồ cautioned writers that they should borrow scientific terms only when necessary.

Incidentally, the Organising Committee of the conference on spelling reforms, knowing that President HỒ had always been concerned over that problem too, sent him a copy of the report. He later received the conference leaders and gave them some of his own ideas about the objectives and methods of spelling reforms as well as his attitude toward the Vietnamese language in general. His comments were published ten years later on the first anniversary of his death in the journal *Ngôn-ngữ*:

Let me remind you to treasure our language. We have to borrow words which we cannot translate from languages of other countries. But we should borrow only when absolutely necessary, and whenever we borrow we must borrow appropriately.

Why do we often borrow unnecessarily and inappropriately? Because (1) we don't treasure our national language and have an inferiority complex, and (2) we have not studied it well and thoroughly.

Borrowing is necessary, but we must fight abuse and laziness. We need a campaign against the overuse of foreign words, the overuse of Chinese loanwords. There are innumerable cases where we can find a Vietnamese word but do not make an effort to find it.

In this respect we must also try to follow the self-help principle in the main, treasure our language, rely on its originality to develop it and use loanwords only as a second resort. We must study our own ways of coining words, we must study more carefully in order to invent our ways of coining new words. (HỒ Chí Minh 1970)

5. The "coining of new words" has been essential in the elaboration of a scientific terminology, which was started even prior to 1945 (Nguyễn Đình-Hoà 1975:47-8). Lê Khả Kế (1968:109-31) gives an account of deliberate steps in terminology work since that date, detailing the use of transliteration, Chinese loanwords and native elements.

An official conference on scientific terminology was convened in Hanoi on December 28 and 29, 1964 by the State Scientific Commission. After a report 'On The Problem of Using Foreign Scientific Terms' by Lưu Văn Lăng and over 20 papers presented by scientists from various sectors, the Conference appointed a committee to continue the study. On May 14 and 21, 1965 the Committee presented its draft proposal of 'Principles of Transliteration of Indo-European Scientific Terms into Vietnamese'. The draft proposal was approved on October 15, 1965 by the Council on Scientific Terminology and Lexicography. In June 1966, the Institute of Social Sciences, nowadays the Social Sciences Commission, recommended that the principles of transliteration be adopted in all social and natural sciences on an *ad hoc* basis. In the brochure devoted to the transliteration of foreign scientific terms into Vietnamese (Ủy-ban Khoa-học xã-hội Việt-nam 1968), it is pointed out that the new

terms must be used creatively, integrated into the Vietnamese language as the technical jargon of the people, thus contributing to the enrichment of the lexical fund. A foreign technical term must be easy to write and to pronounce, yet not too distant from the original form. Each symbol is to represent only one phoneme, and each phoneme is transcribed with only one symbol. The foundation of this consistency is still the speech of the capital city of Hanoi, supplemented by positive elements from other dialects.

Next to transliteration, the source of new scientific and technical terms remains Sino-Vietnamese, that is to say, the extensive Chinese loan-compounds pronounced in the Vietnamese way. A more recent tendency consists in using even Chinese-borrowed forms in the Vietnamese word order, thus preferring a phrase to a compound noun, again if this makes a whole construction clearer: *cụm từ*, not *từ-tổ* 'phrase'.

One example will suffice. Being involved in the late 1950s in the same task within linguistics, this writer together with his colleagues Lê Văn Lý, Nguyễn Khắc Kham and Lê Ngọc Trụ, at the University of Saigon Faculty of Letters tried to coin several needed terms. Although Professor Chao Yuen-ren, in a personal communication, had suggested *ngôn 言* for 'morpheme', I chose *ngữ-vị* on the model of *âm-vị* 'phoneme'.<sup>4</sup> But since my colleague Lê Văn Lý started using it to refer to the European 'morphème', as opposed to *ngữ-nghĩa-vị* 'sémantème', I hesitated for a while between *hình-thái-vị* or *ngữ-thái-vị*, but ended by showing my predilection for the more traditional *từ-tổ*. The Russian-Vietnamese Linguistic Glossary (*Viện Ngôn-ngữ-học* 1969) lists *hình-vị* for 'morpheme' and *từ-vị* for 'lexeme' (pp.44 and 51, respectively).

6. It is tempting to remark that language planning succeeds most through an authoritarian approach. The normative approach, at any rate, is apparent in the case of Vietnam, where the legislator can be either a prescriptive linguist or a literary practitioner.

6.1. Nguyễn Kim Thản (1966:79) cites the 18 pieces of advice which Trường Chinh gave to newspapermen in his book on how to strengthen newspaper writing *Tăng-cường công-tác báo-chí của chúng ta*. All eighteen are 'don't' rules. In order to make the language a national language, newspapermen were told not to use a foreign word unnecessarily, not to write a sentence according to foreign syntax, not to use literary allusions, not to depart from the precious literary tradition of the nation, not to look down on the literary heritage, and not to despise the good elements in foreign literature and arts (p.78). Secondly, in

order to keep the language scientific, newspaper writers were advised not to write a sentence that does not follow Vietnamese grammar, not to use a superfluous word, except for deliberate repetition for emphasis, not to write a sentence which the reader may misunderstand because it is ambiguous, not to make their writing alien to the people's speech, not to write in a disorderly fashion, and not to use old clichés taken either from Vietnamese or from a foreign tongue (although the use of proverbs and sayings would be desirable). The last six rules aimed at making the Vietnamese language closer and closer to the masses. Writers were advised not to be afraid of using ordinary words used by the masses, not to write a sentence which an average reader cannot understand, not to write just for a few elitist intellectuals, not to write lengthy sentences or quote uselessly to impress people, not to neglect the upgrading of the masses' intellectual level just because of eagerness to popularise, and not to use vulgarity just for the sake of making the language accessible to the masses. (p.79).

6.2. Elsewhere a specialist in classical literature enunciates the five principles to be followed to guarantee clarity: (1) Use Vietnamese, stressing the spoken language; (2) Use proverbs, adages, sayings and maxims, which are all extremely precious elements of the literary language of the masses; (3) Coin new words through loan translations from Chinese, thus borrowing the idea and not the phonetic shell; (4) Use Chinese loanwords, since a good writer while being independent does not have to be xenophobic; and (5) Limit the use of Chinese allusions, since the purpose is not to show off, but to make sure everybody understands. (Trương Chính 1972:11).

6.3. The language of President Hồ Chí Minh is among the models upheld because his style is considered simple, concise, to the point and full of imagery. On his 85th birthday, for instance, the College of Pedagogy at Vinh held a seminar May 26-27, 1975 on the late President's skills as a writer.

President Hồ's vocabulary was the subject of a statistical study (Hoàng Cao Cương and Trần Đình Cơ) reported in the journal *Ngôn-ngữ*: this study showed the frequency of nouns (33%), verbs (18%), adjectives (12%), particles (15%) and other lexemes (22%) in a corpus of 58,900 words contained in President Hồ's articles, speeches, reports, appeals and interviews between 1955 and 1959. (Nguyễn Đức Dân, Hoàng Cao Cương and Trần Đình Cơ 1973:14-31).

The manner in which President Hồ explained various socio-political concepts to the masses was discussed in *Cù Đình Tú* (1973:32-6, 62), and his use of idiomatic expressions, sayings and proverbs was studied in *Hoàng Văn Hành* (1973:10-19).

One article (*Bùi Khắc Việt* 1975) was devoted to *Đường Kách Mệnh* 'The Revolutionary Path', a collection of lectures that President Hồ delivered in Kwang-chow in youth training programs in 1926. Although in the 1920s, through short stories and novels by *Phạm Duy Tồn*, *Nguyễn Trọng Thuật*, etc., the Vietnamese language had become streamlined and been moving towards clarity and concision, Hồ Chí Minh's style in this political pamphlet, constructed in the form of questions and answers, was very close to the everyday speech of the people. Carefully chosen vocabulary and simple grammatical structures were supplemented by the use of idioms and vivid imagery taken from native folklore. The book also had a good dosage of rhymes and parallel constructions, which helped take the dryness away from this piece of political discussion, which has been considered "the first foundation of future revolutionary writings." (*Bùi Khắc Việt* 1975:9).

7. At the beginning the agency responsible for language policy implementation was the Linguistics Section of the Institute of Social Sciences. Not having its own publication until 1969 it utilised the medium of *Tap-chí Văn-học*, organ of the Institute of Literature. This review, beginning with Issue No.91 (July 1967), carried a column titled 'Preserve the Purity and Clarity of Vietnamese' (*Giữ-gìn sự trong sáng của tiếng Việt*), and it also promised to print excerpts from a *Handbook of Usage* (*Sổ tay dùng từ*) in progress, whose purpose was to help people study "the standard of modern cultural Vietnamese, i.e., literary Vietnamese - an important and urgent linguistic task." (p.107).

When the Institute of Linguistics, a full-fledged research unit, launched its quarterly journal *Ngôn-ngữ* in 1969, the new publication continued this "usage corner", to which all writers and scholars were invited to contribute. As the organisation with authority to deal with all problems associated with standardisation and purification, the *Viện Ngôn-ngữ-học* has played an important role of co-ordinator and facilitator through its many activities.

During the 5-year period between 1968 and 1972, the Institute of Linguistics accomplished the following (*Ngôn-ngữ* 15 (March 1973):69): (1) compiling a general dictionary and a grammar of Vietnamese; (2) compiling and publishing ten glossaries of social sciences terms and helping with the compilation and publication of another seven volumes



of wordlists in science and technology;<sup>5</sup> and (3) compiling several bilingual dictionaries, English-Vietnamese, French-Vietnamese, Spanish-Vietnamese, Russian-Vietnamese and Chinese-Vietnamese.

Another task no less important has been the study of a policy regarding the languages and writing systems of ethnic minorities in North Vietnam and South Vietnam: it did indeed contribute to the romanisation of Thái, to the transcription of Mường, to the improvement of the Tày-Nùng script and of the Miao script and to the initial investigation of some 47 minority languages of Vietnam. Typical publications in this vital area were a grammar of Tày-Nùng (Hoàng Văn Ma et al 1971), a Mèo-Vietnamese dictionary, a Tày-Nùng-Vietnamese dictionary, a Vietnamese-Mèo dictionary and a Vietnamese-Jarai dictionary. Study of and research on these languages spoken by various nationalities have been steadily pursued since that 1973 report, but the latest account (Hoàng Văn Ma 1975:1-7) makes no mention of the work of the Summer Institute of Linguistics in South Vietnam between 1957 and 1975.

In addition to training courses for in-service cadres and research workers and writings on different topics in grammar and vocabulary, the Institute of Linguistics in Hanoi has issued monographs, organised lectures, seminars and short-term courses, to call people's attention to the problem of preserving the purity and clarity of Vietnamese, and to disseminate essential knowledge about Vietnamese.

8. Vietnamese has assumed every function of a *langue de culture*, serving the national community in every field of endeavour - government, law, diplomatic relations, the press, radio and television, education, science and technology. The Vietnamese lexicon, not only rich but also subtle, has expressive words as well as abstract words and thus can "adequately express every modern idea and feeling", its development being "both quantitative and qualitative" (Nguyễn Kim Thản 1975:6).

As pointed out in Professor Nguyễn Kim Thản's overview (1975:1-14), the policy and guidelines of the Lao-động - now renamed Cộng-sản 'Communist' - Party, on one hand, and President Hồ Chí Minh's model as well as his own contributions to the national language, on the other hand, have shown the serious road to the conscious development and unification of the Vietnamese language. One of the immediate tasks in 1975 was said to be the uniformisation of scientific and technical terminology in each discipline, among various disciplines and particularly between North and South. This problem was considered "more urgent than that of unifying vocabulary in daily life" (p.14). The time had

also come to define the system of phonological standards, which should "be soon used in the schools, over the radio, in TV programs and on the stage." (p.14). At the same time the problem of improving the quốc-ngữ script was said to have to be prepared for the "decisive period".

Undoubtedly, now that the country has been reunified under one political system, government policies regarding language planning and enumerated in this paper will be further implemented so as to make the Vietnamese language a "weapon of the revolution after it was liberated by that revolution throughout three decades." (Nguyễn Kim Thân 1975:14).

In non-dialectic terms the implementation of those well-conceived policies is moving in a centripetal direction, that is, toward the capital city of Hanoi. The following quotes from the same influential scholar are significant:

The basic dialect must be that of the region which is the most developed culturally and which often is also the most developed politically and economically, also the region which has played the greatest part in the historical development of the nation - also the dialect enjoying the most prestige among the people and representing the direction in which the national language is moving. (Nguyễn Kim Thân 1966:14)

The development of literary Vietnamese through several centuries has proved that it is based on the language spoken in the very cradle of the nation and representative of the language of cultural Thăng-Long. We can clearly see this through the poems by Nguyễn Trãi, Đoàn Thị Điểm and down to Nguyễn Du, and from our first works of prose down to the most famous prose texts in Vietnamese literature within the past few decades. We can also see this particularly significant phenomenon whereby the quốc-ngữ script, which had been first used in Saigon, actually did not become widely used among the nation until the Đồng-kinh Nghĩa-thực school in Hanoi disseminated it as the new system of writing. Newspapers and novels, also born in Saigon, did not become firmly rooted in the cultural and intellectual life of the nation and did not positively influence the national language until those products got published in Hanoi, that is to say, in the unified literary language of the entire country. (Nguyễn Kim Thân 1976:12)

If language planning "as a rational and technical process informed by actuarial data and by ongoing feedback is still a dream" (Fishman 1970:111), then in the case of Vietnam, even if it had been a "far-fetched" dream three decades or so ago, parts of that dream are becoming a reality now in late 1976.

N O T E S

1. Neustupny (1970) distinguishes four problems: the code selection problem, involving policy decisions by authorities in control of power; then the second step broken down in two parts, the solution of the problem of the stability of the selected code, requiring codification (by means of dictionaries, grammars, spellers, punctuations, pronunciation guides, etc.), and that of the problem of expansion requiring elaboration (via nomenclatures, thesauruses, etc.). Differentiation involves cultivation through the preparation of style manuals, and the subsidisation of literary creativity in a variety of genres for various purposes and audiences.
2. As in giữ gìn sự trong sáng của tiếng Việt '*preserving the purity and clarity of Vietnamese*' instead of làm cho tiếng ta luôn luôn trong sáng '*make our language always pure and clear*' (p.94).
3. Except in Hanoi all the other schools in North Vietnam in 1966 were still using traditional grammar books (Hồng Dân 1966:58-63).
4. See for instance Nguyễn Đình-Hoà, Bài giảng Ngữ-học Nhập-môn (Saigon: Đại-học Văn-khoa, 1962), 169-86.
5. The bi- and tri-lingual glossaries are listed in Đức Kỳ (1973:31-5).

BIBLIOGRAPHY

BLOOMFIELD, Leonard

1933 *Language*. New York: Holt, Rinehart & Winston.

BUI CÔNG TRÙNG

1961 'Lời bế-mạc hội-nghị'. [Closing Remarks]. In: Viện Văn-học, *Vấn-đề cải-tiến chữ quốc-ngữ* [The Problem of Improving the Quốc-ngữ Script], 382-7.

BÙI KHẮC VIỆT

1975 'Về phong-cách ngôn-ngữ của Bác trong "Đường cách-mệnh"'. [On Uncle Hồ's Linguistic Style in Đường Cách-mệnh]. *Ngôn-ngữ* 24 (June 1975):1-9.

CHAO YUEN-REN

1957 Personal communication.

CHẾ LAN VIÊN

1966 'Làm cho tiếng nói trong sáng, giàu và phát-triển'. [To Make Language Purer, Clearer, Richer and More Developed]. *Tạp-chí Văn-học* 75 (March 1966):28-32.

CÙ ĐÌNH TÚ

1973 'Tìm hiểu cách Hồ Chủ-tịch giảng-giải các khái-niệm cho quần-chúng'. [Understanding the Way President Ho Explained Concepts to the Masses]. *Ngôn-ngữ* 16 (June 1973):32-6, 62.

DƯƠNG LAN HẢI

1972 'Bàn thêm một số điểm xung quanh việc viết hoa tên riêng'. [Some Further Points Regarding the Capitalisation of Personal Names]. *Ngôn-ngữ* 11 (March 1972):53-60.

## ĐẶNG THAI MAI

- 1961 'Thư gửi các bạn trí-thức miền nam'. [Letter to Intellectuals in South Vietnam]. In: Viện Văn-học, *Vấn-đề cải-tiến chữ quốc-ngữ* [The Problem of Improving the Quốc-ngữ Script], 389-92.

## ĐỨC KỲ

- 1973 'Về công-tác biên-soạn từ-điển thuật-ngữ của ta hiện nay'. [On Our Present Task of Compiling Terminological Dictionaries]. *Ngôn-ngữ* 17 (September 1973):31-5.

## FISHMAN, Joshua A.

- 1970 *Sociolinguistics: A Brief Introduction*. Rowley, Mass.: Newbury House.
- 1975 *Language and Nationalism: Two Integrative Essays*. Rowley, Mass.: Newbury House.

## FISHMAN, Joshua A., ed.

- 1974 *Advances in Language Planning*. The Hague: Mouton.

## GARVIN, Paul

- 1959 'The Standard Language Problem - Concepts and Methods'. *Anthropological Linguistics* 1/3:28-31. Reprinted in D. Hymes, ed. *Language in Culture and Society*, 521-6. N.Y.: Harper and Row.
- 1973 'Some Comments on Language Planning'. In: J. Rubin and R. Shuy, eds *Language Planning: Current Issues and Research*, 24-33. Washington, D.C.: Georgetown University Press.

## GARVIN, Paul L. and Madeleine MATHIOT

- 1960 'The Urbanization of the Guaraní Language - A Problem in Language and Culture'. In: Anthony F.C. Wallace, ed. *Men and Cultures: Selected Papers of the Fifth International Congress of Anthropological and Ethnological Sciences*, 783-90. Philadelphia: University of Pennsylvania Press.

## HALL, Robert A. Jr

- 1950 *Leave Your Language Alone!* Ithaca, N.Y.: Linguistica.

## HAUGEN, Einar

- 1966a 'Linguistics and Language Planning'. In: William Bright, ed. *Sociolinguistics*, 50-71. The Hague: Mouton.
- 1966b *Language Conflict and Language Planning: The Case of Modern Norwegian*. Cambridge, Mass.: Harvard University Press.

## HOÀNG PHÊ

- 1961 'Vấn-đề cải-tiến chữ quốc-ngữ'. In: Viện Văn-học, *Vấn-đề cải-tiến chữ quốc-ngữ* [The Problem of Improving the Quốc-ngữ Script], 9-139.
- 1968 'Vấn-đề giữ gìn sự trong sáng của tiếng Việt'. [The Problem of Preserving the Purity and Clarity of Vietnamese]. In: Tổ Ngôn-ngữ-học, *Nghiên-cứu ngôn-ngữ-học* [Studies in Linguistics] vol.1:3-41. Hanoi: Khoa-học Xã-hội.
- 1976 'Một số nguyên-tắc giải-quyết vấn-đề chuẩn-hoá chính-tả'. [Some Principles of the Solution of the Problem of Orthographic Standardisation]. *Ngôn-ngữ* 27 (March 1976): 1-10.

## HOÀNG TUỆ

- 1961 'Tham-luận về vấn-đề cải-tiến chữ quốc-ngữ'. In: Viện Văn-học, *Vấn-đề cải-tiến chữ quốc-ngữ* [The Problem of Improving the Quốc-ngữ Script], 222-34.

## HOÀNG VĂN HÀNH

- 1966 'Tìm hiểu những ý-kiến của Hồ Chủ-tịch về việc mượn và dùng từ gốc Hán'. [Understanding President Ho's Ideas on the Use of Chinese Loanwords]. *Tạp-chí Văn-học* 75 (March 1966): 43-9.
- 1973 'Suy-nghĩ về cách dùng thành-ngữ qua văn-thơ Hồ Chủ-tịch'. [Thoughts on the Use of Idiomatic Expressions in President Ho's Prose and Poetry]. *Ngôn-ngữ* 17 (September 1973):10-19.

## HOÀNG VĂN MA

- 1975 'Tình-hình nghiên-cứu ngôn-ngữ các dân-tộc thiểu-số ở nước ta trong 30 năm qua'. [The Study of Minority Languages in Our Country During the Past 30 Years]. *Ngôn-ngữ* 26 (December 1975):1-7.

HOÀNG VĂN MA, LỤC VĂN PẢO & HOÀNG CHÍ

- 1974 *Từ-điển Tay-Nùng-Việt* [Tay-Nung-Vietnamese Dictionary].  
Hanoi: Khoa-học Xã-hội.

HỒ CHÍ MINH

- 1970 'Một số ý-kiến của Hồ Chủ-tịch về chữ quốc-ngữ và tiếng Việt'. [Some of President Hồ's Ideas on Quốc-ngữ and the Vietnamese Language]. *Ngôn-ngữ* 5 (September 1970):1-3.

HỒ LÊ

- 1972 'Về vấn-đề chuẩn-mực ngôn-ngữ'. [On the Problem of Language Standards]. *Ngôn-ngữ* 11 (March 1972):16-28.

HỒNG DÂN

- 1966 'Vài nét về tình-hình ngôn-ngữ trong nhà trường hiện nay'. [Some Features of the Language Situation in the Schools Today]. *Tạp-chí Văn-học* 75 (March 1966):58-63.
- 1968 'Về quan-hệ giữa mẫu-mực ngôn-ngữ và sự sáng-tạo của cá-nhân'. [On the Relation Between Linguistic Norms and Individual Creativity]. In: *Tổ Ngôn-ngữ-học, Nghiên-cứu ngôn-ngữ-học* [Studies in Linguistics], vol.1:53-67. Hanoi: Khoa-học Xã-hội.

JERNUDD, Björn H. and Jyotirindra DAS GUPTA

- 1971 'Towards a Theory of Language Planning'. In: Rubin and Jernudd, eds 1971:195-215.

LÊ KHẢ KẾ

- 1968 'Elaboration of a Vietnamese Scientific Terminology'. In: *Vietnamese and Teaching in Vietnamese in D.R.V.N. Universities*, 109-31. Hanoi: Foreign Languages Publishing House.

LÊ NGỌC TRỤ

- 1973 *Việt-ngữ Chánh-tả Tự-vị* [Dictionary of Vietnamese Orthography]. Revised edition of 1961 edition.

LÊ VĂN THỚI & NGUYỄN VĂN DƯƠNG

- 1968 'Nguyên-tắc soạn-thảo danh-từ chuyên-khoa'. [Principles of Coining Technical Terms]. *Danh-từ Chuyên-môn* [Scientific Terminology] 1 (September 1968):3-23. Saigon: Ủy-ban Quốc-gia Soạn-thảo Danh-từ Chuyên-môn.

## LÊ XUÂN THẠİ

- 1966 *Ngôn-ngữ là gì?* [What is Language?]. Hanoi: Khoa-học.
- 1968 'Mẫu-mực và phát-triển'. [Norms and Development]. In: *Tổ Ngôn-ngữ-học, Nghiên-cứu ngôn-ngữ-học* [Studies in Linguistics], vol.1:42-52. Hanoi: Khoa-học Xã-hội.

## LƯU VÂN LĂNG, NGUYỄN KIM THẦN &amp; NGUYỄN VĂN TÚ

- 1961 *Khái-luận ngôn-ngữ-học* [Elements of Linguistics]. Hanoi: Trường Đại-học Tổng-hợp Hà-nội. Quoted in Hoàng Phê 1968.

## MARTINET, André

- 1963 *Éléments de linguistique générale*. 3rd edition. Paris: Armand Colin.

## NEUSTUPNY, Jiri B.

- 1970 'Basic Types of Treatment of Language Problems'. *Linguistic Communications* 1:77-100.

## NGUYỄN BĂNG HỒ

- 1966 'Vài ý-kiến nhỏ về cách phát-âm trên sân khấu kịch nói'. [Some Small Ideas About the Pronunciation on the Stage]. *Tạp-chí Văn-học* 76 (April 1966):72-5.

## NGUYỄN ĐĂNG LIÊM

- 1975 'Cases and Verbs in Pidgin French (Tây bồi) in Vietnam'. In: Herman H. van Olphen, ed. *Linguistic Borrowing, Working Papers, 1974 Conference, American Council of Teachers of Uncommonly-taught Asian Languages, Occasional Papers, No.2*: 15-35. Austin, Texas: Center for Asian Studies.

## NGUYỄN ĐÌNH-HOÀ

- 1955 *Quốc-ngữ: The Modern Writing System of Vietnam*. Washington, D.C.
- 1960 'Toward the Standardization of Vietnamese Orthography'. *Journal of the Burma Research Society* 43/1 (June 1960):67-72.
- 1962 *Bài giảng Ngữ-học Nhập-môn* [Lectures on Introductory Linguistics], Tập 1. Saigon: Đại-học Văn-khoa.



- 1966 Review of A Vietnamese Grammar by Laurence C. Thompson.  
*Language Learning* 16/3-4:213-28.
- 1975 'Lexical and Syntactic Borrowing in Modern Vietnamese'.  
In: Herman H. van Olphen, ed. *Linguistic Borrowing, Working Papers, 1974 Conference, American Council of Teachers of Uncommonly-taught Asian Languages*, Occasional Papers, No.2: 36-59. Austin, Texas: Center for Asian Studies.
- NGUYỄN ĐỨC DÂN, HOÀNG CAO CƯƠNG & TRẦN ĐÌNH CƠ
- 1973 'Bước đầu tìm hiểu sự phân-bố từ-vựng trong ngôn-ngữ Hồ Chủ-tịch'. [Preliminary Research on Lexical Distribution in President Ho's Language]. *Ngôn-ngữ* 16 (June 1973):14-31.
- NGUYỄN HÀM DƯƠNG
- 1974 'Một số vấn-đề về tiêu-chuẩn-hoá tiếng Việt'. [Some Problems Concerning the Standardisation of Vietnamese]. *Tổ-quốc* 330 (March 1974):13-16.
- 1975 'Mấy vấn-đề về chuẩn-hoá tiếng Việt'. [Some Problems Concerning the Standardisation of Vietnamese]. *Ngôn-ngữ* 23 (January 1975):26-34.
- NGUYỄN KHÁNH TOÀN
- 1973 'Bước ngoặt vĩ-đại và ngành ngữ-học'. [The Great Turning Point and Linguistics]. *Ngôn-ngữ* 15 (March 1973):1-4.
- NGUYỄN KHẮC-KHAM
- 1976 'Vietnamese National Language and Modern Vietnamese Literature'. *East Asian Cultural Studies* 15/1-4 (March 1976):177-94.
- NGUYỄN KIM THẢN
- 1961 'Tham-luận về vấn-đề cải-tiến quốc-ngữ'. In: Viện Văn-học, *Vấn-đề cải-tiến chữ quốc-ngữ* [The Problem of Improving the Quoc-ngu Script], 199-221.
- 1966 *Rèn luyện về ngôn-ngữ* [Training in Language]. Hanoi: Khoa-học.

- 1974 'Tiếng Việt trên con đường tiến lên của dân-tộc'.  
[Vietnamese on the Forward Road of the Nation]. *Tổ-quốc*  
330 (March 1974):11-13.
- 1975 'Thử nhìn lại chặng đường lịch-sử của tiếng Việt trong 30  
năm qua'. [Let us Look Back at the Historic Journey of  
the Vietnamese Language During the Past Thirty Years].  
*Ngôn-ngữ* 25 (September 1975):1-14.
- 1976 'Tính thống-nhất của tiếng Việt'. [The Unity of the  
Vietnamese Language]. *Văn-hoá Nghệ-thuật* 56 (April 1976):  
10-12.
- NGUYỄN KIM THẦN, HỒ LÊ, LÊ XUÂN THẠI & HỒNG DÂN
- 1967 *Nói và viết đúng tiếng Việt* [Speak and Write Vietnamese  
Correctly]. Hanoi: Khoa-học Xã-hội.
- NGUYỄN TUÂN
- 1966 'Về tiếng ta'. [About Our Language]. *Tạp-chí Văn-học* 75  
(March 1966):21-7.
- NGUYỄN VĂN HUYỀN
- 1966 'Nhà trường và thầy giáo có nhiệm-vụ chăm lo giữ gìn sự  
trong sáng của tiếng Việt'. [Schools and Teachers have  
the Task of Preserving the Purity and Clarity of the  
Vietnamese Language]. *Tạp-chí Văn-học* 75 (March 1966):6-8,  
95-6.
- NGUYỄN VĂN THẠC
- 1968 'Vấn-đề lạm-dụng từ Hán-Việt'. [The Excessive Use of Sino-  
Vietnamese Words]. In: *Tổ Ngôn-ngữ-học, Nghiên-cứu ngôn-  
ngữ-học* [Studies in Linguistics], vol.1:106-22. Hanoi:  
Khoa-học Xã-hội.
- NOSS, Richard
- 1967 *Language Policy and Higher Education. Higher Education  
and Development in Southeast Asia 3/2*. Paris: UNESCO and  
The International Association of Universities.
- OLPHEN, Herman H. van, ed.
- 1975 *Linguistic Borrowing. Working Papers, 1974 Conference,  
American Council of Teachers of Uncommonly-taught Asian*

*Languages*. Occasional Papers, 2. Austin, Texas: Center for Asian Studies.

PHẠM VĂN ĐỒNG

- 1966 'Giữ gìn sự trong sáng của tiếng Việt'. [Preserving the Purity and Clarity of Vietnamese]. *Tạp-chí Văn-học* 75 (March 1966):1-5, 93-5. Also in *Học-tập* 123 (April 1966): 7-13.

PHAN THIỀU

- 1972 'Bàn về qui-tắc viết hoa tên người, tên đất trong tiếng Việt'. [Principles of Capitalisation of Personal Names and Place Names in Vietnamese]. *Ngôn-ngữ* 11 (March 1972): 61-6.

QUANG DẠM

- 1966 'Vấn-đề giữ gìn sự trong sáng của tiếng Việt đối với những người viết báo Việt-Nam'. [The Problem of Preserving the Purity and Clarity of Vietnamese as it Concerns Vietnamese Newspapermen]. *Tạp-chí Văn-học* 76 (April 1966):60-71, 91-8.

RAY, Punya Sloka

- 1963 *Language Standardization: Studies in Prescriptive Linguistics*. The Hague: Mouton.

RUBIN, Joan and Björn H. JERNUDD, eds

- 1971 *Can Language Be Planned? Sociolinguistic Theory and Practice for Developing Nations*. Honolulu: East-West Center and University of Hawaii Press.

RUBIN, Joan and Roger SHUY, eds

- 1973 *Language Planning: Current Issues and Research*. Washington, D.C.: Georgetown University Press.

SẮC-LỆNH SỐ 125/VHXX NGÀY

- 8-7-1966 [Decree of July 8, 1966 establishing the Institute of Culture]. Saigon.

SAUSSURE, Ferdinand de

- 1960 *Cours de linguistique générale*. 5th edition. Paris: Payot.

T. Tr.

- 1948 'Hãy gây một phong-trào Việt-hoá lời nói và văn-chương'.  
[Let us Create a Movement to Vietnamise both Language and Literature]. *Sự thật*, December 19, 1948.

THANH HÀ

- 1968 'The Languages of National Minorities and the Creation or Improvement of Their Scripts'. In: *Mountain Regions and National Minorities in the D.R. of Viet-Nam*. Vietnamese Studies, 15:121-36. Hanoi: Vietnamese Studies.

THOMPSON, Laurence C. and David D. THOMAS

- 1967 'Vietnam'. In: Thomas A. Sebeok, ed. *Current Trends in Linguistics*, vol.2: *Linguistics in East Asia and South East Asia*, 815-46. The Hague: Mouton.

TỔ NGÔN-NGŨ-HỌC

- 1968 *Nghiên-cứu ngôn-ngữ-học. Tập 1, Vấn-đề giữ gìn sự trong sáng của tiếng Việt* [Studies in Linguistics. Vol.1: The Problem of Preserving the Purity and Clarity of the Vietnamese Language]. Hanoi: Khoa-học Xã-hội.

TRƯƠNG CHÍNH

- 1972 'Cha ông ta đã phấn-đấu như thế nào để ngôn-ngữ văn-học dân-tộc ngày càng trong sáng và phong-phú'. [How Our Forefathers strove to make the National Literary Language become increasingly purer and richer]. *Ngôn-ngữ* 12 (June 1972):1-11, 20.

TRƯỜNG CHÍNH

- 1951 *Tăng-cường công-tác báo-chí của chúng ta* [To Strengthen Our Newspaper Writing]; quoted by Nguyễn Kim Thân 1966: 78-9.

ỦY-BAN KHOA-HOC XÃ-HỘI. VIỆT-NAM

- 1968 *Qui-tắc phiên thuật-ngữ khoa-học nước ngoài ra tiếng Việt* [Principles of Transliteration of Foreign Scientific Terms into Vietnamese]. Hanoi: Khoa-học Xã-hội.

VIỆN NGÔN-NGŨ-HỌC

- 1969 *Thuật-ngữ ngôn-ngữ-học Nga-Việt* [Linguistic Terminology: Russian-Vietnamese]. Hanoi: Khoa-học Xã-hội.

- 1972 'Dự-thảo qui-tắc viết hoa'. [Draft on the Principles of Capitalisation]. *Ngôn-ngữ* 11 (March 1972):49-52.

## VIỆN VĂN-HỌC

- 1961 *Vấn-đề cải-tiến chữ quốc-ngữ* [The Problem of Improving the Quốc-ngữ Script]. Hanoi: Viện Văn-học.

## VIỆT-NAM THÔNG TẤN XÃ

- 1972 *Vấn-đề sử-dụng nhân-danh và địa-danh ngoại-ngữ* [The Use of Personal Names and Place Names from Foreign Languages]. Saigon: VNNTX. No.7869, October 5, 1972, morning edition, pp.K1-K9.

## VŨ BÁ HÙNG

- 1966 'Một số vấn-đề về ngôn-ngữ trong phim truyện của ta hiện nay'. [Some Language Problems in our Present Feature Movies]. *Tạp-chí Văn-học* 76 (April 1966):75-80.

## VƯƠNG LỘC

- 1975 'Coup d'oeil sur l'évolution de la langue vietnamienne'. *Essais Linguistiques*. Etudes Vietnamiennes 40:9-30. Hanoi: Etudes Vietnamiennes.

## XUÂN DIỆU

- 1966 'Sự trong sáng của tiếng Việt trong thơ'. [The Purity of Vietnamese in Poetry]. *Tạp-chí Văn-học* 75 (March 1966): 9-20, 27.



# A STATISTICAL COMPARISON OF VERBS AND NOUNS IN ROGLAI

MAXWELL COBBEY

Word counts were made of four Roglai texts,<sup>1</sup> distinguishing between (1) noun-words, (2) verb-words, and (3) all other words. All words were classified contextually as to grammatical class, so that, for instance, noun-words included adjectives, classifiers, quantifiers, etc., in addition to nouns and pronouns, whenever they functioned endocentrically with nouns, or as noun substitutes. However, in another place an adjective might be counted as a verb-word because of its verbal function in that context.

## 1. WORD AND CLUSTER COUNTS

Besides this word count of grammatical types, a count was also made of the number of clusters of each of the three types, a cluster being defined as a consecutive string of one or more words of the same grammatical type. These counts of word types and cluster types are given in Table I.

TABLE I  
Word and Cluster Counts

Name of Text	Word Count				Cluster Count			
	Nouns	Verbs	Other	Total	Nouns	Verbs	Other	Total
History	1512	753	394	2659	603	601	305	1509
Monkey	156	137	57	350	92	91	42	225
Eagle	124	78	45	247	68	42	42	152
Feast	325	209	66	600	173	153	63	389
<b>Totals</b>	<b>2117</b>	<b>1177</b>	<b>562</b>	<b>3856</b>	<b>936</b>	<b>887</b>	<b>452</b>	<b>2275</b>

By making calculations based on the counts listed in Table I, various comparisons were made (a) between the four Røglai texts, and (b) between the grammatical types.

## 2. RELATIVE NUMBER OF WORDS WITHIN TEXTS<sup>2</sup>

Noun-words are more numerous than any other type in all the texts; and two of the texts each consist of more than half noun-words, as seen in Table II. I feel that these two texts, History and Feast, are by far the two best texts from a literary standpoint.

TABLE II  
Relative Number of Words within Texts

Text	Noun	Verb	Other	Total
History	0.57	0.28	0.15	1.00
Monkey	0.45	0.39	0.16	1.00
Eagle	0.50	0.32	0.18	1.00
Feast	0.54	0.35	0.11	1.00
Average <sup>3</sup>	0.55	0.31	0.15	1.00

## 3. RELATIVE NUMBER OF CLUSTERS WITHIN TEXTS

A comparison of the number of clusters of each grammatical type within the texts (Table III) shows that for two of the texts, the History and the Monkey, there are about the same number of noun and verb clusters. These two texts contain more action, whereas the Eagle and the Feast are more descriptive, and consequently show more noun clusters than verb clusters.

TABLE III  
Relative Number of Clusters within Texts

Text	Noun	Verb	Other	Total
History	0.40	0.40	0.20	1.00
Monkey	0.41	0.40	0.19	1.00
Eagle	0.45	0.28	0.28	1.00
Feast	0.45	0.39	0.16	1.00
Average <sup>3</sup>	0.41	0.39	0.20	1.00



#### 4. CLUSTER LENGTHS

From Table IV below it is seen that noun clusters are longer on the average than verb clusters. The two best literary texts (History and Feast) are seen to have the greatest difference between noun cluster length and verb cluster length.

TABLE IV  
Words per Cluster

Text	Noun Cluster	Verb Cluster	Other Cluster	Average
History	2.51	1.25	1.29	1.76
Monkey	1.70	1.51	1.36	1.56
Eagle	1.82	1.86	1.07	1.62
Feast	1.88	1.37	1.04	1.54
Average <sup>3</sup>	2.26	1.33	1.24	1.69

#### 5. RELATIVE CLUSTER LENGTHS WITHIN TEXTS

The relative cluster length of each grammatical type within a text can be seen in Table V, which is obtained by dividing the cluster lengths (of Table IV) by the average cluster length of a text (the last column of Table IV). Here it is seen that the two best literary texts have the widest range of relative cluster length between grammatical types.

TABLE V  
Relative Cluster Lengths within Texts

Text	Noun	Verb	Other	Range <sup>4</sup>
History	1.42	0.71	0.73	0.71
Monkey	1.09	0.97	0.87	0.22
Eagle	1.12	1.14	0.66	0.48
Feast	1.22	0.89	0.68	0.54
Average <sup>3</sup>	1.33	0.78	0.73	0.60

#### 6. RELATIVE CLUSTER LENGTHS BETWEEN TEXTS

A comparison of cluster lengths between texts for a given grammatical type was made by dividing the cluster length of each text by the average cluster length for the type, as in Table VI. Verbs are seen to have the widest range of relative cluster lengths between the texts.

TABLE VI  
Relative Cluster Length Between Texts

Text	Noun	Verb	Other	Average <sup>3</sup>
History	1.11	0.94	1.04	1.04
Monkey	0.75	1.13	1.09	0.92
Eagle	0.81	1.40	0.86	0.96
Feast	0.83	1.03	0.84	0.91
Range <sup>4</sup>	0.36	0.46	0.25	0.13

## 7. DENSITY OF GRAMMATICAL TYPES

An empirical density coefficient was calculated for each cluster type within each text. This was obtained by multiplying the entries of Table V by the entries of Table II. The greater the relative cluster length of a grammatical type, the greater its coefficient of density. And the greater the relative number of words of that type in a text, the greater the coefficient of density.<sup>5</sup>

TABLE VII  
Density of Grammatical Types within Texts

Text	Noun	Verb	Other
History	0.81	0.20	0.09
Monkey	0.49	0.38	0.14
Eagle	0.56	0.36	0.12
Feast	0.66	0.31	0.07
Average <sup>3</sup>	0.73	0.24	0.11

## 8. LITERARINESS

As mentioned previously, the two best literary texts have the greatest differences between average noun cluster length and verb cluster length. Another characteristic of some good literary texts is relative shortness of clusters, due to the grammatical types being intermixed in semi-poetical types of sentences. These two somewhat opposing characteristics of cluster length are combined in an empirical formula for literariness, Table VIII, which gives the absolute difference between noun and verb cluster length divided by the average length of clusters in that text. The History and Feast texts come out with the highest coefficients in literariness.

TABLE VIII  
Coefficient of Literariness

History	0.71
Monkey	0.12
Eagle	0.02
Feast	0.33
Word-weighted average	0.53
Non-weighted average	0.30

### 9. A SAMPLE ENGLISH COMPARISON

Similar counts and calculations were made of an article in English for comparison, using an article from the *Reader's Digest*, October 1975: 73-7.

	Noun	Verb	Other	Total
I. Word counts	1033	346	344	1723
Cluster counts	537	198	285	1020
II. Rel. no. of words	0.60	0.20	0.20	
III. Rel. no. of clusters	0.53	0.19	0.28	
IV. Cluster length	1.92	1.75	1.21	
V. Rel. cluster length	1.14	1.03	0.71	
VI. (only one text)				
VII. Density of gram. types	0.68	0.21	0.14	
VIII. Literariness coeff.				0.10

From II and III above, the English text is seen to have a greater relative number of both noun words and of noun clusters than the Rŏglai texts. And the English text has a smaller relative number of verb words and verb clusters than any of the Rŏglai texts.

Tables IV and V reveal that both the cluster length and the relative cluster length of the English text fall within the range of the Rŏglai texts. The noun and verb density coefficients (VII) of the English text also fall within the range of the Rŏglai texts.

The coefficient of 'literariness' of the English text is lower than three of the Rŏglai texts. We defined 'literariness' solely for Rŏglai as having short clusters and a great difference between noun and verb cluster length. Good literary style in English clearly has different characteristics from Rŏglai.

## N O T E S

1. Røglai is a language belonging to the Coastal Chamic branch of Austronesian, found in south Vietnam inland from Nhatrang to Phanthiet.
2. The relative number of nouns in the History text, for instance, is found by dividing the number of nouns in the History text, 1512 (Table I), by the total number of words in the History text, 2659 (Table I): which gives 0.57.
3. These averages are not obtained by averaging the figures given in this chart, but must be obtained from the original data of Table I; for instance, for the nouns of Table II  $2117 \div 3856 = 0.55$ .
4. The range, for instance, of relative cluster length in the History text is  $1.42 - 0.71 = 0.71$ ; for the Monkey text  $1.09 - 0.87 = 0.22$ .
5. This may also be stated: within each text the density coefficients of the respective grammatical types vary in direct proportion to both the relative lengths of clusters of each type and the relative numbers of words of each type. It may be calculated directly for each grammatical type by the formula:

$$\text{Density coefficient}_A = \frac{(\text{Words}_A)^2}{(\text{Words}_T)^2} \times \frac{\text{Clusters}_T}{\text{Clusters}_A}$$

where A = the given grammatical type

T = the total words (or clusters) of all types.

# TWO PROGRESSIVE ASPECT MARKERS IN CHINESE

ANNA KWAN-TERRY

## 1. THE USE OF zài

This article sets out to show that there are two progressive aspect markers (PAM) in Chinese which exhibit certain similarities as well as differences in meaning. The two markers are zài 在 and -zhe 着. The meaning of zài will be examined first. Compare the following sentences:

tāmen	yǐjīng	zài	hē	tāng	le
他们	已经	在	喝	汤	了
they	already	(PAM)	drink	soup	(marker of new situation)

'They are already taking their soup.'

tāmen	yǐjīng	hēle	tāng	le
他们	已经	喝了	汤	了
they	already	drink (perfect aspect marker)	soup	(marker of new situation)

'They have already taken their soup.'

tāmen	yǐjīng	hēguò	tāng	le
他们	已经	喝过	汤	了
they	already	drink (experiential aspect marker)	soup	(marker of new situation)

'They have already taken their soup.' OR

'They have taken soup before.'

tāmen	yǐjīng	yào	hē	tāng	le
他们	已经	要	喝	汤	了
they	already	(imminent aspect marker)	drink	soup	(marker of new situation)

*'They are already about to take their soup.'*

These sentences are distinguished from one another in that the action of taking soup is seen to be at different phases of development. In the first sentence, the action is in progress; in the second, it is completed; in the third, it is completed and forms part of the experience of the persons referred to; and in the last, it is imminent or about to take place. Next, compare the following:

háizi	yòu	zài	kū	le
孩子	又	在	哭	了
child	again	(PAM)	cry	(marker of new situation)

*'The child is again crying.'*

háizi	yòu	kūqǐlái	le
孩子	又	哭起来	了
child	again	cry (inchoative aspect marker)	(marker of new situation)

*'The child again starts crying.'*

In the first sentence, the action is in progress whereas in the second it is inchoative. Then too contrast the meaning of *zài*, *-le* (to be distinguished from *le*, the marker of new situation, which is not bound to a preceding verb and which occurs finally in a clause), *-guò*, *yào* and *-qǐlái* in the above sentences and that of *-xiàqù* in the following sentence:

bié	kūxiàqù	le
别	哭下去	了
don't	cry (continuative aspect marker)	(marker of new situation)

*'Don't keep on crying.'*

Here *-xiàqù* indicates the continuation of the action of crying from a specific time. When *zài* is thus seen to be in contrast with *yào*, *-le*, *-guò*, *-qǐlái* and *-xiàqù* and to form with them (as well as with *-zhe* which will be discussed later) a closed system of items all indicating phase of development or aspect, there is reason to believe that *zài* is a grammatical aspect marker on a par with the others. This is further supported by the observation that although *zài* (as well as *yào*)

is not morphologically bound to the verb as -le, -guò, -qǐlái and -xiàqù (which are generally recognised as grammatical aspect markers of the language) are, it is never used independently of the verb; on the contrary it is always tied to it, with or without intervening adverbial elements. In this light, zài (and yào) may justifiably be considered an affix, not to the verb itself, but to the verb phrase (which may consist of the verb alone or the verb plus adverbial elements). For these two reasons, it is proposed here that zài (as well as yào) should be recognised as a grammatical aspect marker of Chinese.

It has been suggested earlier that zài is used to indicate an action in progress. It is necessary to examine the meaning of this aspect marker in greater detail. Examples of the use of this marker are useful for this purpose.

- (1) chēhùo fāshēng de shíhòu, tāmen zài shuōxiào  
 车祸 发生 的 时候, 他们 在 说笑  
 car accident happen (marker of subordination) time, they (PAM) joke

*'When the car accident happened, they were joking.'*

- (2) dāng wǒ líkāi de shíhòu, tāmen zài tāolùn  
当 我 离开 的 时候, 他们 在 讨论  
when I leave (marker of sub- time, they (PAM) discuss  
ordination)

zhèjiàn	shì
这件	事
this (noun	matter
classifier)	

*'When I left, they were discussing this matter.'*

- (3) wǒmen      zhěnggè      xiàwǔ      dōu      zài      niàn shū  
我们      整个      下午      都      在      唸书  
we      whole (noun      afternoon      all      (PAM)      study  
classifier)

*'We were studying the whole afternoon.'*

- (4) jiǎshǐ míngtiān wǒ huílái de shíhòu nǐmen  
假使 明天 我 回来 的 时候 你们  
if tomorrow I return (marker of sub-ordination) time you

hái	zài	dǎ	pái,	wǒ	biàn	kāichú	nǐmen
还	在	打	牌,	我	便	开除	你们
still	(PAM)	play	card,	I	consequently	dismiss	you

'If you are still playing cards when I return tomorrow, I will dismiss you.'

- (5) wúlùn wǒ shénme shíhòu huí jiā, tā dōu zài  
 无论 我 什么 时候 回家, 他都 在  
 no matter I what time return home, he always (PAM)  
 kàn shū  
 看 书  
 read book  
 'No matter what time I return home, he is always reading.'

- (6) kàn! nàzhī mǎ zài chī cǎo  
 看! 那只 马 在 吃 草  
 look! that (noun classifier) horse (PAM) eat grass  
 'Look! that horse is eating grass.'

- (7) wàimiàn zài xià - yǔ  
 外面 在 下雨  
 outside (PAM) come down rain  
 'It's raining outside.'

- (8) shāngkǒu zài chū - xuè  
 伤口 在 出 血  
 wound (PAM) come out blood  
 'The wound is bleeding.'

In all the above sentences, a specific time is erected to which the action referred to in the sentence is related as being in progress at the time. In sentences (1) to (5), this time is explicitly stated: the time of the car accident, the time of the speaker's departure, a specific afternoon, the time when the speaker returns the next day, and whenever the speaker returns home. In sentence (6), this time juncture is understood to be the time when the speaker points with his finger at the horse and utters the sentence. This time juncture however need not be explicitly referred to in a sentence but may be understood from the context of discourse. The point to be noted is that whenever *zài* is used (the same applies to other aspect markers), a specific time is either explicitly or implicitly referred to. In sentences (7) and (8), in the absence of evidence pointing to the contrary, the time could be said to be the time of utterance.

Since *zài* indicates action in progress and since an action can only be said to be in progress if it takes up an appreciable period of time, a corollary to this is that *zài* carries a certain durative quality. It is this durative quality in *zài* which probably accounts for the incompatibility between *zài* and such verbs as *dàodá* 到达 'arrive',



huílái 回来 'return', yíshī 遗失 'lose', fāxiàn 发现 'discover', and sǐ 死 'die' that belong to an aspectual-semantic group of verbs which may be labelled transitional event verbs. One semantic characteristic of these verbs is that the transition designated by these verbs is momentary (hence the incompatibility between these verbs and zài), though the approach to that transition may take up an appreciable period of time.

In discussing the meaning of zài, it is useful to make reference to a distinction Comrie draws between the imperfective and the perfective (the term perfective is not to be confused with the term perfect: the latter refers to the phase when an action has been completed. So too the term imperfective is not to be confused with the term progressive, though some have used the former to refer to the progressive aspect). A perfective verb, according to Comrie (1976:3-4):

... presents the totality of the situation referred to without reference to its internal temporal constituency: the whole of the situation is presented as a single unanalysable whole, with beginning, middle and end rolled into one; no attempt is made to divide this situation up into the various individual phases that make up the action of entry. Verb forms with this meaning will be said to have perfective meaning, and where the language in question has special verbal forms to indicate this, we shall say that it has perfective aspect .... The other forms ... do not present the situation in this way, but rather make explicit reference to the internal temporal constituency of the situation.

The other forms are the imperfective verb forms. In the light of this distinction, a verb marked by zài can be said to be an imperfective verb form for it presents the action as stretched out in time so that the action is analysable into various internal phases and attention is drawn to the middle of the action, as it were. Verbs marked by -qǐlái and -xiàqù are imperfective in this sense as they indicate the beginning and the continuation of an action respectively. Verbs marked by -le, -guò and yào, on the other hand, are perfective verbs as they present an action as a single whole, the whole of which is completed in the case of -le and -guò and the whole of which is about to occur in the case of yào.

It should be noted that zài is used only when the linguistic or situational context calls for the adoption of the particular angle of vision or aspect emphasis that zài carries. In making a timeless statement like

niú gēng tián, mǎ chī cǎo  
 牛 耕 田, 马 吃 草  
 cow work field, horse eat grass

'Cows work the field and horses eat grass.'

For example, there is no reason for seeing the actions as stretched out in time, nor any reason for putting emphasis on a specific time juncture when the actions are in their course of development, so *zài* is not used. Instead, the actions are seen as whole units without any aspect implications and so the unmarked verb form is used. Similarly, if a child wants to find out what a particular horse generally eats, without reference to any particular time juncture, he says

māma, nàzhī mǎ chī shénme?  
 妈, 那只 马 吃 什么?  
 mother, that (noun classifier) horse eat what

'Mother, what does that horse eat?'

However, if he wants to find out what that horse is eating at a specific time, then he says

māma, nàzhī mǎ zài chī shénme?  
 妈, 那只 马 在 吃 什么?  
 mother, that (noun classifier) horse (PAM) eat what

'Mother, what is that horse eating?'

or

māma, nàzhī mǎ chīzhe shénme?  
 妈, 那只 马 吃着 什么?  
 mother, that (noun classifier) horse eat (PAM) what

'Mother, what is that horse eating?'

The use of *-zhe* will be discussed at a later stage. A speaker will also choose the unmarked verb form if he wishes to indicate the occurrence of an action (whether realised or potential) without focusing attention on any specific time and the stage of development the action has reached at this time. So he says:

gāngcái tāmen yíge tán qín, yíge  
 刚才 他们 一个 弹 琴, 一个  
 just now they one (noun classifier) play piano one (noun classifier)

chàng gē  
 唱 歌  
 sing song

'Just now one of them played the piano and one of them sang.'

Here each of the two actions is presented as a single whole like a dot on a line representing the time continuum. The time reference indicates the location of this dot on the line (the time-when of the action) and not an independently selected time to which the action is related. The same can be said of

děng	yíhuǐr	nǐmen	yíge	tán	qín,	yíge
等	一会儿	你们	一个	弹	琴	一个
wait	a while	you	one (noun classifier)	play	piano	one (noun classifier)

chàng	gē
唱	歌
sing	song

'In a little while, one of you play the piano and one of you sing.'

where each of the two actions is seen as a single whole to be performed in the future. In the earlier sentence with past time reference as well as in this second sentence with future time reference, the actions performed or to be performed are seen as whole units and there is no call for looking into their internal temporal constituency and singling out one particular phase within their course of development for special mention. It is the absence of any need to call attention to the actions as in their course of development that accounts for not using *zài* in both sentences.

## 2. THE USE OF -zhe

Apart from *zài*, there is another marker which is used to indicate action in progress at a specific time. This is the marker -zhe. Scholars of the Chinese language have unanimously agreed that it is an aspect marker of the language. Some have suggested that it may be used interchangeably with *zài*, though they prefer to classify the latter as an adverb rather than a grammatical aspect marker. Zhāng Zhì-gōng (1957:94), for example, suggests that

... when a certain action, event or state is seen at a given point of time to be in progress or in a state of continuance, then generally the adverb zhèng 正, zài 在 or zhèng zài 正在, etc. is added before the verb, or the verb suffix -zhe is added after it, or both may be used simultaneously. [the translation is mine]

The same point has earlier been made by Gāo Míng-kǎi (1948:377) who says: "When an occurrence is seen as in continuous development ... this is expressed in the spoken language by -zhe, zài, zhèng zài ... -zhe." [the translation is also mine].

An important point in the present article is that *zài* and *-zhe* are distinguished in meaning and use, though both involve the notion of something (either an action or a state) going on or in progress. The term 'in progress' here is used in a very broad sense to include, not only the progress or development of an action (this is the sense in which the term is generally used), but also the continuance of a state. In other words, the term 'in progress' is here used stripped of its dynamic implication. This particular use of the term is suggested by the use of *-zhe* in Chinese which reflects that a sentence indicating an action in progress is not necessarily dynamic in nature. A sentence like

tā	shǒulǐ	názhe	yībēn	shū
他	手裡	拿着	一本	书
he	in hand	carry (PAM)	one (noun classifier)	book

*'He is carrying a book in his hand.'*

can hardly be said to be dynamic. To accommodate this meaning and use of *-zhe*, the term 'in progress' is used in the present study in a broad sense stripped of any inherent dynamic implication.

*-zhe* is used in a number of ways. The first use is exemplified in the following sentences:

zhuōzishàng	bǎizhe	yì	píng	huā
桌子上	擺着	一	瓶	花
table	place (PAM)	one	vase	flower

*'On the table is placed a vase of flowers.'*

qiángshàng	guàzhe	yìfú	huà
牆上	掛着	一幅	画
wall	hang (PAM)	one (noun classifier)	picture

*'On the wall is hung a picture.'*

shìnèi	fàngzhe	liǎngtiáo	bǎn	dòng
室內	放着	兩條	板	凳
room	place (PAM)	two (noun classifier)	wood	bench

*'In the room are placed two wooden benches.'*

chuángshàng	duīzhe	hěn	duō	yángwáwa
床上	堆着	很	多	洋娃娃
bed	heap (PAM)	very	many	doll

*'Many dolls are heaped on the bed.'*

which all have the structure

locative expression + verb with -zhe + noun phrase

With sentences of this kind, the action designated by the verb is not itself in progress; instead, it is completed. The effect of the action however is sustained at the time in question. What is 'in progress' or in continuance, therefore, is not the action but the state resulting from it. This interpretation is supported by the observation which Fàn Fāng-lián (1963:386-95) has made that the -zhe in such sentences can be replaced by the perfect aspect marker -le, without affecting the meaning of the sentences. The first of the above sentences can be used to illustrate the point: when

zhuōzishàng	bǎizhe	yì	píng	huā
桌子上	擺着	一	瓶	花
table	place (PAM)	one	vase	flower

is changed to

zhuōzishàng	bǎile	yì	píng	huā
桌子上	擺了	一	瓶	花
table	place (perfect aspect marker)	one	vase	flower

there is no change in meaning at all. In the last sentence, the use of -le indicates that the action of placing the vase of flowers on the table is completed. But the completion of the action also means that a new state prevails, the state of the vase of flowers being on the table. Here it can be seen that the distinction between a dynamic predication and a stative predication is not always maintained. Here both sentences are essentially expressive of a state despite the fact that in both an action is indicated. This stative quality of the two sentences under discussion can be supported by the fact that the verb in both can be replaced by either the verb of existence *yǒu* 有 or the verb 'to be' *shì* 是 without affecting the meaning of the sentences, as Fàn Fāng-lián (1963) has pointed out. Instead of the two sentences where the verb is marked by -zhe or -le, one can equally say

zhuōzishàng	$\left[ \begin{array}{c} yǒu \\ 有 \\ exist \\ shì \\ 是 \\ be \end{array} \right]$	yì	píng	huā
桌子上		一	瓶	花
table		one	vase	flower

'On the table is a vase of flowers.'

without introducing any essential change in meaning.

It is interesting to note that in sentences with the structure

locative expression + verb with -zhe + noun phrase

zài cannot be used in place of -zhe: it is impossible to have

*zhuōzishàng	zài	bǎi	yì	píng	huā
桌子上	在	擺	一	瓶	花
table	(PAM)	place	one	vase	flower

This points to one essential difference between zài and -zhe: whereas a verb (or verb phrase) marked by zài is necessarily accompanied in a sentence by a noun (or noun phrase) which represents the agent responsible for the action, -zhe is not restricted in use in this way.

In a sentence like

zhuōzishàng	bǎizhe	yì	píng	huā
桌子上	擺着	一	瓶	花
table	place (PAM)	one	vase	flower

*'On the table is placed a vase of flowers.'*

neither of the two noun phrases, zhuōzishàng and yì píng huā, represents the agent. The use of zài in place of -zhe in such a sentence, as has been said, is impossible. Where zài is used, the sentence carries obligatorily a noun phrase representing the agent, as in

tā	zài	tán	qín
他	在	彈	琴
he	(PAM)	play	piano

*'He is playing the piano.'*

and

dìqiú	zài	xuánzhuǎn
地球	在	旋轉
earth	(PAM)	revolve

*'The earth is revolving.'*

-zhe is used to indicate the continuance of a state resulting from the completion of an action in a different kind of sentence. Here the agent is given, as in

(1) tā	chuānzhe	yíjiàn	xīn	yīfù
她	穿着	一件	新	衣服
she	wear (PAM)	one (noun classifier)	new	dress

*'She is wearing a new dress.'*

- (2) tā dòng yě bú dòng de zhànzhe  
 他 动 也 不 动 地 站着  
 he move also not move (marker of sub-ordination) stand (PAM)

'He stands motionless.'

- (3) tā zhěngtiān dōu guānzhe mén  
 他 整天 都 关着 门  
 he all day all close (PAM) door

'He keeps the door closed all day.'

- (4) tā chángcháng bìzhe yǎnjīng  
 他 常常 闭着 眼睛  
 he often shut (PAM) eye

'He often keeps his eyes shut.'

One possible interpretation of sentence (1) is that the action of putting on a new dress is completed but the effect persists: the person referred to is still wearing her new dress. In sentence (2), the action of standing up is completed but the effect of someone being in a standing position is still in continuance. In sentence (3), the action of closing the door is completed, but the resultant state of the door being closed prevails. In sentence (4), the action of shutting one's eyes is completed, but the resultant state of having one's eyes shut still persists. Here it can be seen how sometimes the same verb-base is used to indicate an action which brings into being a certain state as well as to indicate the resultant state of the action: so *chuān yīfù* 穿衣服 'put on one's clothes', *zhànqǐlái* 站起来 'stand up' (note here that the addition of the directional complement *qǐlái* is necessary to indicate the action of standing up as distinguished from the state of being in a standing position, *zhànzhe* 站着), *guān mén* 关门 'close the door' and *bì yǎnjīng* 闭眼睛 'shut one's eyes' bring in respectively the states of *chuānzhe yīfù* 穿着衣服 'having one's clothes on', *zhànzhe* 站着 'being in a standing position', *guānzhe mén* 关着门 'having one's door closed' and *bìzhe yǎnjīng* 闭着眼睛 'having one's eyes shut'. In all these instances in which -zhe is used, -zhe indicates that the state resulting from the performance of an action is sustained (the state is in progress, as it were, to use the term in a broad sense).

A somewhat similar use of -zhe is its use with what are now currently referred to as stative verbs or adjectives as they are traditionally called. In the following sentences, -zhe indicates that a certain state is sustained or in progress:

tā de yǎnjīng zhǒngzhe, yīfù hái  
 他 的 眼睛 肿着, 衣服 还  
 he (marker of sub- eye swollen (PAM), clothes still  
 ordination)

shīzhe  
 湿着  
 wet (PAM)

'His eyes are swollen and his clothes are still wet.'

háizi de biànzǐ-shāo quǎnzhe, yòu  
 孩子 的 辫子梢 卷着, 又  
 child (marker of sub- pig-tail end curl (PAM), also  
 ordination)

huáng yòu jiāo  
 黄 又 焦  
 yellow also burnt

'The tips of the child's pig-tails are curled, (looking) both yellowish and burnt.'

tiān hái zǎozhe, tàiyáng hái xiézhē  
 天 还 早着, 太阳 还 斜着  
 day still early (PAM), sun still slant (PAM)

'The day is still early and the sun is still in a slanting position.'

dà fēng bǎ liǔ-shù chuīde wānzhe  
 大 风 把 柳树 吹得 弯着  
 strong wind (marker of object willow tree blow till bend (PAM)  
 pre-position)

'The strong wind blew till the willow tree was bent.'

tā hóngzhe liǎn, yíjù huà yě bù shuō  
 他 红着 脸, 一句 话 也 不 说  
 he red (PAM) face, one (noun word also not say  
 classifier)

'He reddened his face (blushed) and didn't say a word.'

the last sentence may also be written as

tā de liǎn hóngzhe, yíjù huà  
 他 的 脸 红着, 一句 话  
 he (marker of sub- face red (PAM), one (noun word  
 ordination) classifier)

yě bù shuō  
 也 不 说  
 also not say

'His face was red and he didn't say a word.'



In these sentences, the state that is sustained is sometimes clearly the result of some initial transitional event, so a willow tree is in a bending position because it has been bent; in other cases, as in *tiān hái zǎozhe*, it is difficult to say whether any transitional event is involved. In all such sentences in which the verb is a stative verb and as such partakes of a stative or descriptive quality, the use of *zài* in place of *-zhe* is totally unacceptable. This points to the conflict between *zài* and a stative predication.

Another use of *-zhe* is exemplified in the following:

wàimiàn	chuīzhe	dà	fēng
外面	吹着	大	风
outside	blow (PAM)	strong	wind

*'Strong winds are blowing outside.'*

shìnnèi	màozhe	yì-céng-céng	yānwù
室内	冒着	一层层	烟雾
room	come out (PAM)	layers	smoke

*'Smoke is coming out from the room.'*

tā	zhěngtiān	dāidāi	de	wàngzhe	xiàngpiàn
他	整天	呆呆	地	望着	相片
he	all day	blank	(marker of subordination)	stare (PAM)	photo

*'He keeps staring at the photo blankly all day.'*

fùqīn	duì	wǒ	dèngzheyǎn,	xiàomàzhe
父亲	对	我	瞪着眼,	笑着
father	at	me	open wide (PAM) eye,	joke scold (PAM)

*'Father kept his eyes wide open at me and kept scolding me jokingly.'*

where the actions are clearly seen as in progress. Here there is no question of the actions being completed. However, in this use of *-zhe* to indicate action in progress, *-zhe* seems to carry a certain stative quality. This stative quality is most apparent in sentences where *-zhe* is used to mark a subordinate verb in a verb-plus-verb series, as in

tā	xiàozhe	shuō
他	笑着	说
he	laugh (PAM)	speak

*'He spoke smilingly.'*

tā pǎozhe lái  
他 跑着 来  
he run (PAM) come  
'He came running.'

tā suǒzhe méitóu xiǎng zhègè wèntí  
他 锁着 眉头 想 这个 问题  
he knit (PAM) brow consider this (noun classifier) question  
'Knitting his brows, he considered this question.'

In all three cases, the first verb marked by -zhe is subordinated to the second, main verb and describes the action which accompanies the second action. In other words, the first verb describes the manner in which the action designated by the main verb is done. This subordinate verb marked by -zhe therefore has a descriptive function and reflects the stative quality in this verb form. In the following sentence:

wǒ mànmande zǒu, shùzhe dìshàng de  
我 慢 慢 地 走, 数 着 地 上 的  
I slowly walk, count (PAM) ground (marker of subordination)

hóng zhuān  
红 砖  
red brick

'I walked slowly, counting the red bricks on the ground.'

the verb marked by -zhe designates the action which accompanies the action of walking. Here again the verb marked by -zhe takes on a descriptive or stative quality, being expressive of the manner in which another action is done. A verb in the -zhe form may also appear as the main verb in a sentence, as in

nà tiān wǎnshàng, tāmen liǎng rén zài hǎibīn  
那 天 晚上, 他们 两 人 在 海滨  
that day night, they two persons in seaside  
màn bùzhe, mèngxiǎngzhe měimǎn de jiāng lái  
漫步着, 梦想着 美 满 的 将来  
stroll (PAM), dream (PAM) beautiful (marker of subordination) future

'That night, the two of them strolled (were strolling) along the seaside, and dreamed (were dreaming) of their beautiful future.'

where a scene, rather than an action as such, is depicted. In prose writing in particular, a sentence where the main verb is marked by -zhe is used more for background description than for plot-advancing narration. This again points to the descriptive quality in -zhe.

It is worth noting that in all instances in which a verb marked by -zhe is subordinated to another verb and represents the action which accompanies the main action, the use of zài in place of -zhe is impossible. It is not possible, for example, to have

*tā	zài	suǒ	méitóu	xiǎng	zhègè	wèntí
他	在	锁	眉头	想	这个	问题
he	(PAM)	knit	brow	consider	this (noun classifier)	question

or

*wǒ	mǎnmǎnde	zǒu,	zài	shù	dìshàng	de
我	慢 慢 地	走,	在	数	地上	的
I	slowly	walk,	(PAM)	count	ground	(marker of subordination)

hóng	zhuān
红	砖
red	brick

The impossibility of such a substitution supports the theory that whereas a verb marked by -zhe carries a certain descriptive or stative quality and leads one to see an action in progress as a state of action sustained, a verb marked by zai is more dynamic and leads one to see an action in progress as a dynamic carrying out of an activity rather than as a state of action sustained.

### 3. CO-OCCURRENCE OF zài AND -zhe WITH SEMANTIC VERB GROUPS

One way of highlighting the similarity as well as difference between these two aspect markers is to show their possibilities of co-occurrence with various semantic verb groups. It appears that while there are certain semantic verb groups which co-occur with both aspect markers, there are others which are repugnant to them, and yet others which co-occur with -zhe but not with zài, thereby reflecting an essential difference between the two markers.

The first group of verbs may be termed dynamic durative verbs and are identifiable by their compatibility with both zài and -zhe. Semantically these verbs indicate actions which may be extended over a period of time, and examples include:

kū	哭	'cry'
xiào	笑	'laugh'
tán-huà	谈话	'talk'
chuī-kǒushào	吹口哨	'whistle'

xǐ-zǎo	洗澡	'bathe'
biáoyǎn	表演	'perform'
xiǎng	想	'think'

Within this group of verbs there is a sub-group each member of which designates an action which, taken singly, is instantaneous and yet each is infinitely repeatable and may therefore be performed over an extended period of time in the form of a series of repeated acts. This sub-group includes verbs like:

hūxī	呼吸	'breathe'
késòu	咳嗽	'cough'
tī	踢	'kick'
tiào-dòng	跳动	'jump up and down'

and when these verbs co-occur with *zài* or *-zhe*, a series of repeated acts rather than one single occurrence of an action is meant.

On the other hand, as has been said, there are verb groups which resist both *zài* and *-zhe*. One such group includes members like:

shì	是	'be'
yǒu	有	'have, exist'
děngyú	等於	'equal'
bāokuò	包括	'include'
shǔyú	屬於	'belong to'
zhí	值	'be worth'
dàibiǎo	代表	'represent'

for which the term relation verbs may be used. These verbs are recognisable by being either totally repugnant to the perfect aspect marker *-le* or, if they take this marker, the marked and the unmarked forms are identical in meaning. For example, *bāokuò* and *bāokuòle* are not distinguishable in meaning. (It has to be pointed out that in this group of relation verbs, there is one particular member, *yǒu*, which is found to co-occur with *-zhe* under certain limited and describable circumstances, as Zōu Guó-tǒng (1956:39) has pointed out. However, where *-zhe* is used with *yǒu*, *-zhe* is void in meaning and may be dropped without affecting the meaning of the sentence in any way.)

The group of verbs denoting various psychic states like

míngbái	明白	'understand'
xiāngxìn	相信	'believe'
xìnrèn	信任	'trust'
zhīdào	知道	'know'
pèifú	佩服	'admire'

are also repugnant to both *zài* and *-zhe*. It is interesting to note that, like the relation verbs, these psychic verbs either resist the perfect aspect marker *-le*, or if they take this marker, the marked and the unmarked forms are not distinguished in meaning.

The group of auxiliary verbs with members like

kéyǐ	可以	'can'
huì	会	'will probably'
yīnggāi	应该	'should'
gǎn	敢	'dare'

are also incompatible with *zài* and *-zhe*. Indeed, a defining feature of these verbs is their inability to take any aspect markers at all. Besides the semantic verb groups just mentioned, there are also certain verbs or verb phrases which are either aspectual verbs in themselves in that they designate certain aspectual notions like

kāishǐ	开始	'begin'
jìxù	继续	'continue'
wán	完	'finish'

or they carry some aspectual particle, as

kàn-jiàn	看見	'see'
yù-dào	遇到	'meet'
kǎo-shàng	考上	'take examination and pass it'
chī-wán	吃完	'finish eating'

These verbs or verb phrases which inherently convey some aspectual notion are found to resist co-occurrence with *zài* and *-zhe*.

Then finally there are certain verb groups which, while they resist co-occurrence with *zài*, are found to take *-zhe*. One such verb group may be referred to semantically as verbs of existence, including items like

cúnzài	存在	'exist'
shēngcún	生存	'exist'
huó	活	'live'
mímàn	弥漫	'pervade'

Since existence is not a dynamic action, the use of zài with these verbs is impossible. However, the state of being in existence may be sustained and hence -zhe is found to co-occur with these verbs. Another group of verbs which resist zài but may co-occur with -zhe is the group of what may be referred to as posture verbs, like

zhàn	站	'stand'
zuò	坐	'sit'
tǎng	躺	'lie'
kǎo	靠	'lean against'

When these verbs are marked by -zhe, they indicate the result of the performance of some initial action designated by the same verb-base (with or without a directional complement). So the action of zhànle qǐlái 站了起来 'having stood up' results in the state of zhànle zhān 站着 'standing', and the action of kǎole zài yǐzishàng 靠在椅子上 'having leaned against the chair' results in the state of kǎozhe yǐzi 靠着椅子 'leaning against the chair'. One last group of verbs need to be mentioned: this is the group of stative verbs or adjectives which are totally repugnant to zài but some of which are used with -zhe to indicate the continuance of a state, as in

wāizhe zuǐ-jǎo	歪着咀角	'keeps one's mouth corners in a slanting position'
qīngzhe bùfá	轻着步伐	'keeps one's steps soft'
hēizhe liǎn	黑着脸	'keeps one's face dark' (i.e. gloomy-looking)
chénmòzhe	沉默着	'keeps silent'
tiándì fāngzhe	田地荒着	'field remains unattended to'

In summing up, it can be said that both zài and -zhe are used to indicate an action or state in continuance or in progress, using the latter term 'in progress' in a broad sense to include continuance of a state. For this reason, they may both be termed progressive aspect markers. This also explains why these two aspect markers are sometimes found to co-occur with each other, though each resists co-occurrence

with other aspect markers in the system. However, there is also an important difference between zài and -zhe which is that the former is much more dynamic in nature than the latter: zài is used to indicate an action in active progress, whereas -zhe is used to indicate the continuance of a state which may be a state resulting from the completion of an action or simply a state of being which does not essentially involve an active action, or it is used to lead one to see an action in progress as a state of action sustained. This difference makes it impossible for one to be used in place of the other in certain circumstances.

BIBLIOGRAPHY

COMRIE, Bernard

- 1976      *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems.* Cambridge University Press.

FÀN FĀNG-LIÁN

- 1963      'cúnzài jù'. *Zhōngguó yǔwén.*

GĀO MÍNG-KǎI

- 1948      *hànyǔ yǔfǎ lùn.* Kāimíng shūdiàn.

ZHĀNG ZHÌ-GŌNG

- 1957      *hànyǔ yǔfǎ chángshì.* Xīnzhīshí Chūbǎnshè.

JIŌU GUÓ-TŌNG

- 1956      'tántán 'yǒuzhě''. *Zhōngguó yǔwén.*



## A NOTE ON THE TAGALOG PASSIVE IN THE TOTANES MANUSCRIPT

JOSEPH F. KESS

It is interesting to note that Fray Sebastian de Totanes (1865) treats the Tagalog passive in a manner highly predictive of the later treatment of the Tagalog verbal paradigm in the descriptive linguistic period. Totanes' *Arte de la lengua tagala* was one of the earlier<sup>1</sup> and most influential descriptions of the Tagalog language by the early Spanish grammarians. Its influence can be seen as reaching beyond the Spanish grammatical period itself and reflections of Totanes' insights and approaches can be directly seen in Blake 1906, 1916, 1925, and, to a lesser degree, Bloomfield 1917.

Totanes devotes the entire second section known as the *libro segundo*<sup>2</sup> to the treatment of various aspects of the verb and participle. In Chapter II of the *libro segundo* Totanes provides twelve rules to aid in the use and understanding of the verbs in Tagalog. Of particular interest are rules five (*de la formacion de las pasivas*), six (*del uso de las tres pasivas*), and seven (*cuando se ha de hablar por activa, y cuando por pasiva*).

Totanes makes the same observation regarding the passives that is to be found in the later works cited above by Blake and Bloomfield, as well as an observation that is replicated in more recent descriptive works. According to Totanes, it would appear that "en este idioma es mas frecuente el hablar por pasiva que por activa" (p.28). This is undoubtedly still the case, with the so-called passives seemingly outnumbering the so-called actives. But what has never been completely sorted out is how to predict the active-passive split in voice-marking verbal constructions for given Tagalog verbs, nor which apparent voice-marking affix will predict which apparent focus construction relationship in a sentence with that given verb. Kess (1975, and 1979) has attempted to point out

the fact of multiple overlap and surface ambiguity attached to some affix uses and the difficulty of expecting certain verbal construction types to automatically and unambiguously account for certain semantic focus types. In fact, it is obvious that this one inescapable fact has probably done more than any other factor to enhance the acceptance of semantic investigations as probably holding the key to the focus, topic and case relationships in the Tagalog verbal paradigm.

Secondly, the only reason that such an observation about passives outnumbering actives might even be made is for the historical reason that the first studies of this Austronesian language type is that they were carried out by linguists and grammarians working from the framework of Indo-European linguistic structures. Indeed, here the active is undoubtedly more prevalent than the corresponding passive; documentation to support this position can be readily found in recent psycholinguistic investigations in English and other languages (for example, see Goldman-Eisler and Cohen 1970). Had the situation been reversed historically, the observation would likely also have been reversed. But what is more interesting is that this fact negates the need or use for preserving the active as any kind of starting point in descriptions of Tagalog. There may, however, be some merit to using one particular sentential construction type as a starting point for pedagogical purposes, as was the case with the Bowen text (see Bowen 1965). Here, of course, the question begins to make for divergent approaches - should one take the most numerous construction type as the base form type for descriptive purposes or should one take as base form type the construction type from which other form types can be most easily derived. The question is obviously one with pragmatic overtones on both sides, but one is pragmatism as a reflection of psycholinguistic reality and the other a reflection of the so-called simplicity metric in linguistic practices.

At any rate, this returns one to the basic question of when the passives appear or a given passive appears. Nothing much in this respect has changed since Totanes' time; as he puts it, "una de las mayores dificultades de esta lengua, es saber usar de estas pasivas, segun su significacion" (p.29). The very fact of multiple overlap between the use of some of the case-marking affixes made for difficulties in the structuralist period because form and form alone could not be trusted to always provide unambiguous clues to both syntactic construction type or semantic focus type. This same problem arises for Totanes, for his definitions are based on logico-semantic designations of the sentence types. For example, the definition provided for verbal constructions in general struggles to capture what generality there is, and in at

least one case, the passive in *i-* (*y-* for Totanes) must perforce be fairly broad in its scope. Thus, as Totanes notes,

La regla general es esta: todo lo que es echar, ó como echar hacia fuera, ó apartar de sí: todo lo que es instrumento, ó que tiene veces de él: que en castellano se dice con; todo lo que es causa, ó como causa, respeto, reverencia, ó motivo; que en castellano se dice por, y todo lo que es, determinado tiempo pide pasiva de y.

The definitional statements for the other passives are no clearer, though they are shorter in this respect. For example, in the case of *-an* and *-in-*, Totanes makes the following observation:

Todo lo que es lugar, ó como lugar, pide la pasiva de an.  
Y todo lo demas que no es esto pide la pasiva de in.

As Totanes himself admits, the rule is far from fool-proof, but serves as a general indicator of what different semantic intentions seem to be encompassed by the several verbal constructions. As Totanes himself puts it,

Esta regla bien entendida era suficiente, para saber cuando se ha de usar una pasiva, y cuando de otra; pero respecto de la dificultad, es muy general esta regla, y se necesita de alguna individuacion en cada una de estas pasivas ....

This, however, is exactly the point, and little seems to have changed since the publication of the Totanes manuscript. The key to the formulation of the so-called passives in Tagalog is not one that can rest exclusively on form alone, but is of necessity one that depends upon the semantic intent underlying the sentence construction. The point of the commentary here presented is that only with the attention shift to matters semantic does one seem to entertain any hope of seeing the function and deployment of the Tagalog verbal paradigm revealed. Totanes himself foreshadows modern semantically-oriented studies like those investigations by Naylor (1973) and Ramos (1974). He did note, and was probably the first published grammarian-linguist to do so for Tagalog, that the key may lie in exactly this area of inquiry. For example, Totanes observes that:

Es precision el hablar por activa, siempre que se habla de cosa indeterminada; lo que se conocerá en no llevar alguno de los artículos, los, las, le, de los, de las, etc., ni derivativo, meus, tuus, vester, etc., ni demostrativo alguno como este, ese, de aquel, aquello, etc., que son los determinantes. Vg.: Mata una gallina, es indeterminado porque no dice cual gallina, etc. Y así precisa á hacerse por activa; matay ca nang isang manúc. Trae agua, lo mismo: Magdalá cà nang tubig .... Llama un muchacho. Tomauag ca nang isang bata.

Por lo opuesto, precisa el hablar por pasiva, siempre que llevase la oracion alguno de aquellos determinantes de la

cosa. Vg.: Mata la gallina. Patain mo ang manuc. Trae aquella agua. Dalhin mo yaong tubig .... Llama á mi muchacho. Tauagin mo ang aquing batà .... Lo mismo se dice de otra caulquiera que tenga esta u otra equivalencia.

As a concluding remark, one may observe that it is indeed interesting that little seems to have been resolved on the topic of the exact deployment of the Tagalog verbal paradigm during the early structuralist period, for the simple reason that form is not as trustworthy as it was often taken to be. On the other hand, there was ample suggestion from the pen of Totanes himself that the key may lie in the semantic content of notions like focus and topic, rather than tying it to structural considerations. The Totanes *Arte de la lengua tagala* stands thus as both a valuable initial repository of vital suggestions in terms of avenues of research as well as an object lesson in how the rubrics of one period in the history of scientific concerns are translated into the working methodologies of another period. The moral of the story is that something is always lost in the translation.

N O T E S

1. The reference used here as well as most commonly cited for this manuscript is for the 1865 edition of the *Arte de la lengua tagala*. For example, see Constantino 1971. There were three preceding editions of the manuscript: 1745 (Sampaloc), 1796, and 1850. My thanks are due to Mrs Alice Loranth, Curator of the John G. White collection of Orientalia in the Cleveland Public Library, for briefly making one of the earlier editions (1796) available for comparison.
2. The original manuscript appears in a Spanish slightly archaic by modern standards, and its use of diacritics (e.g. accent placement) differs from that of the modern orthographic practice. Since Spanish is probably a language that followers of scholars of Philippine studies are likely familiar with, excerpts are left in the original to avoid missing the appropriate nuance of Totanes' turn of phrase.

BIBLIOGRAPHY

ELAKE, Frank R.

- 1906 'Expression of Case by the Verb in Tagalog'. *JAOS* 27:183-9.
- 1916 'The Tagalog Verb'. *JAOS* 36:396-414.
- 1925 *A Grammar of the Tagalog Language*. American Oriental Series  
1. New Haven, Conn.: American Oriental Society.

BLOOMFIELD, Leonard

- 1917 *Tagalog Texts with Grammatical Analysis*. University of  
Illinois Studies in Language and Literature 3/3.

BOWEN, J.D., ed.

- 1965 *Beginning Tagalog: A Course for Speakers of English*.  
Berkeley: University of California Press.

CONSTANTINO, Ernesto

- 1971 'Tagalog and Other Major Languages of the Philippines'.  
In: T.A. Sebeok, ed. *Current Trends in Linguistics*, vol.8:  
*Linguistics in Oceania*, 112-54. The Hague: Mouton.

GOLDMAN-EISLER, F. and M. COHEN

- 1970 'Is N,P, and PN Difficulty a Valid Criterion of Trans-  
formational Operations?'. *Journal of Verbal Learning and  
Verbal Behavior* 9:161-6.

KESS, Joseph F.

- 1975 'On the Semantics of Focus'. *Anthropological Linguistics*  
17/7:353-62.

- 1979 'Focus, Topic and Case in the Philippine Verbal Paradigm'.  
In: Nguyen Dang Liem, ed. *South-East Asian Linguistic Studies*, vol.3. *Pacific Linguistics*, C-45:211-37.
- NAYLOR, Paz Buenaventura  
1973 Topic, Focus and Emphasis in the Tagalog Verbal Clause.  
Doctoral dissertation, University of Michigan, Ann Arbor.
- RAMOS, T.V.  
1974 *The Case System of Tagalog Verbs*. PL, B-27.
- TOTANES, S. de  
1865 *Arte de la lengua tagala*. 4th edition. Binondo: Miguel Sanchez y Ca.





## KELEY-I PHONOLOGY AND MORPHOPHONEMICS

LOU HOHULIN and MICHAEL KENSTOWICZ

Keley-i is a Malayo-Polynesian language spoken by approximately 5,000 people on Central Luzon in the Philippines between the Ibaloi and Ifugao territories. Although culturally Keley-i speakers are Ifugao, their language shares features of both adjacent families and has recently been placed in a separate group called Kallahan by Reid (1975). This study is based on the fieldwork of the first author and her husband who visited the Keley-i area many times between 1965 and 1974.

There are 23 phonemes in Keley-i, 18 consonants and five vowels.

p	t	c	k	ʔ	i	u
b	d	ɟ	g		e	o
	s	ʃ		h		a
m	n		ŋ			
w	l	y				

The palatal obstruents are inordinately rare, having arisen from palatalisations of dentals by y or occurring in borrowings from neighbouring languages where these sounds are as common as the stop consonants.<sup>1</sup> Aside from borrowed words like ʔiskul 'school' and ʔistet 'States', s is a variant of t before i, as can be seen by the following forms where the infix -in- has been placed inside roots beginning with t: takang 'to open the mouth', s-in-ekang; tugun 'to advise', s-in-ugun. The remaining consonants have relatively free distribution except for n, which assimilates to the point of articulation of a following consonant, and ʔ. The glottal stop alternates with Ø in two situations. First, stems ending in a vowel take an inserted ʔ before pause and before a vowel. Thus, gesi 'to carve' appears as gesiʔ at the end of a phrase

and as *gesi*ʔ-an when the suffix -an has been appended to the stem. Although /*gesi*ʔ/ could be viewed as the underlying form and a rule deleting ʔ before words beginning with a consonant invoked, the fact that u glides to w before a vowel<sup>2</sup> militates against this approach: cf. *bayu* 'to pound rice' which appears as *bayu*ʔ before pause but as *bayw*-an before a vowel. If /*bayu*ʔ/ were the underlying form, an otherwise unnecessary and unnatural rule deleting ʔ in the context u\_\_V would be required. On the other hand, if the underlying forms are /*bayu*/ and /*gesi*/, all that is required is to order the rule gliding prevocalic u before the rule that inserts ʔ between two vowels. The other ʔ-∅ alternation occurs with the affixes /in/ and /um/ which can appear either prefixed or infix. Their underlying shape emerges when they are infix, while when prefixed a glottal stop is inserted since no Keley-i word begins with a vowel phonetically: cf. the forms of *dilag* 'to light' d-in-ilag, d-um-ilag, ʔin-dilag, ʔum-dilag. It should be noted that all roots which begin with a glottal stop must be assumed to have that sound present in their underlying forms, since the glottal stop is not lost, even when preceded by a consonant-final prefix: cf. ʔinum 'to drink', man-ʔinum. The contexts in which the glottal stop is inserted in Keley-i are summarised in the following rule.

$$\emptyset \rightarrow ʔ / \left\{ \begin{array}{l} \# \_ V \\ V \_ V \\ V \_ \text{pause} \end{array} \right\}$$

Turning to the vowels, e is always lax while i and u are lax only before syllable-final velars. Aside from a few words like *to*ʔon 'year', o is limited to borrowings.

As in other languages of the Philippines, much of the morphology of the verb in Keley-i is connected with the highlighting or bringing into focus of a particular NP in the sentence. The syntactic aspects of this Philippine phenomenon have been treated by a number of writers, most recently by Schachter (1976). Since we are concerned with the phonology and morphology of the Keley-i verb, only a brief description of the syntactic facts will be presented here.

In every sentence a particular NP is highlighted or made the 'topic'. It will usually correspond to a NP marked by the definite article in English. Non-focused NPs in the sentence are usually indefinite. Pronouns occur in a focus and a non-focus set (also oblique) while nouns are marked by a binary system of particles.

	pronoun sets			particles	
	focus	non-focus		focus	non-focus
1sg.	ʔak	ku	personal	hi	nan hi
2sg.	ka	mu	non-personal	hu/∅	ni
3sg.	∅	tu	locative	di/∅	di
ldu.	ʔita	ta			
1pl.excl.	kami	mi			
1pl.incl.	itsu	tayu			
2pl.	kayu	yu			
3pl.	ʔida	da			

The verb is marked by a portmanteau morpheme expressing tense and 'agreement' with the highlighted NP in terms of its syntactic function in the sentence. Thus, compare the two sentences below.

- (a) b-imm-edbed hi Juan ni paʔul.  
*'John has bound some cane.'*
- (b) b-in-edbed nan Juan (hu) paʔul.  
*'John has bound the cane.'*

In (a) the subject 'John' is in focus, while the object 'cane' (paʔul) is not. This is marked by the appropriate particles. The past tense on the verb bedbed 'to bind' is marked by the infix -imm- since it is the subject of the sentence that is highlighted. On the other hand, in (b) paʔul is in focus and 'John' is not. This necessitates a different agreement marker on the verb. In this case the past tense is marked by -in-, used to signal object focus. It should be noted that the word order in Keley-i is fixed: Verb-Subject-Object. Transposition of the subject and object in either (a) or (b) leads to an ungrammatical sentence. Keley-i seems to differ in this respect from some other Philippine languages where the word order is apparently more free.

In addition to subject and object focus, Keley-i has what we will call (c) accessory focus (used for instrumentals), (d) referent focus (used when the action of the verb is directed to an object located in a particular region of time or space), and (e) beneficial focus. Examples of the past tense of bedbed in each of these foci follow.

- (c) ʔim-bedbed nan Juan ʔetan ʔikat ni paʔul.  
   that string  
*'John has bound the cane with that string.'*
- (d) b-in-edbed-an nan Juan hu heli tu.  
   leg his  
*'John has bound (it, e.g. the wound) on his leg.'*

(e) ʔim-bedbed-an nan Juan hi Pablo ni paʔul.

*'John has bound some cane for Paul.'*

It should be pointed out that a Keley-i sentence frequently has only two NPs per clause and for this reason some of the above sentences, while perfectly grammatical, are a bit unnatural. When there are more than two NPs in a clause, the focus NP must appear in one of the first two NP slots. The subject must appear immediately after the verb, even if it is not in focus. The position of the object, referent, accessory, and beneficial NPs is dependent on which of these NPs is in focus. This constraint accounts for the fact that the object paʔul occupies the third NP slot in (c) and (e).

Most of the morphophonemic alternations in Keley-i occur in the verb inflection. In each of the five simple foci enumerated above, a verb appears in an imperative form and three tense forms: past, present, and future. The basic opposition is between past and non-past or more accurately between perfective (completed action) and imperfective (incompleted action). The imperfective is broken down into present (action begun) vs. future (action not yet begun). As we shall see, these two basic oppositions are reflected in various ways in the morphology. The overwhelming majority of verb roots are disyllabic, occurring in two canonical forms: CVCV(C) or CVCCV(C). Examples of the latter type are ʔagtu *'to carry on the head'* and duntuk *'to punch'*.

subject focus	fut.	man-ʔagtu	ʔum-duntuk
	past	nan-ʔagtu	d-imm-untuk
	pres.	ka-man-ʔagtu	ka-ʔum-duntuk
object focus	fut.	ʔagtu-ʔen	duntuk-en
	past	ʔ-in-agtu	d-in-untuk
	pres.	ka-ʔagtu-ʔa	d-in-untuk
accessory focus	fut.	ʔi-ʔʔagtu	ʔi-dduntuk
	past	ʔin-ʔagtu	ʔin-duntuk
	pres.	ke-ʔi-ʔʔagtu	ke-ʔi-dduntuk
referent focus	fut.	ʔagtu-ʔan	duntuk-an
	past	ʔ-in-agtu-ʔan	d-in-utuk-an
	pres.	ka-ʔagtu-ʔi	ka-duntuk-i
beneficial focus	fut.	ʔi-ʔʔagtu-ʔan	ʔi-dduntuk-an
	past	ʔin-ʔagtu-ʔan	ʔin-duntuk-an
	pres.	ke-ʔi-ʔʔagtu-ʔi	ke-ʔi-dduntuk-i

Inspection of these paradigms reveals that, by and large, the present and future are differentiated from the past in the same way and that

the present is differentiated from the future by the prefix *ka-*. Although further analysis of the inflectional morphemes is possible (e.g. the beneficial seems to combine the prefixes of the accessory with the suffixes of the referent focus) we will ignore this here. Two morphophonemic rules are operative in these paradigms. First, the *a* of the present morpheme *ka-* is raised to *e* when followed by an *i*. The vowel *e* also triggers this change (cf. *ke-bedbed-an* 'bind' obj. f. pres.). Other prefixes of the shape *Ca-* do not exhibit this behaviour and so this raising rule is limited to just this one morpheme. Second, note that the root-initial consonants are geminated after the prefix *ʔi-*. This rule, like a number of other consonant gemination rules in Keley-i, is limited to the present and future forms of the verb and thus supports the opposition between perfective and imperfective. For example, the imperative in the accessory focus is also marked by *ʔi-*: cf. *ʔi-galgal* 'chew!' vs. *ʔi-ggalgal* acces. future; *ʔi-bedbed* 'bind!' vs. *ʔi-bbedbed* acces. future. Also, in stative forms the prefix *ʔi-* appears in all tenses. Here we find gemination only in the present and future, as seen below.

stative	fut.	me-ʔi-ggalgal-an	me-ʔi-bbedbed-an
	past	ne-ʔi-galgal-an	ne-ʔi-bedbed-an
	pres.	ke-ʔi-ggalgal-an	ke-ʔi-bbedbed-an

We thus require a rule geminating the root initial consonant after the prefix *ʔi-* in the imperfective aspect.

We now turn to roots of the shape CVCV(C), which are much more susceptible to morphophonemic change. The simplest types are those whose first root vowel is high. For example, *pili* 'to choose' and *duyag* 'to pour' are inflected as follows.

	subj. f.	obj. f.	acces. f.	ref. f.	ben. f.
fut.	ʔum-pilli	pilli-ʔen	ʔi-ppili	pilli-ʔan	ʔi-ppili-ʔan
past	p-imm-ili	p-in-ili	ʔim-pili	p-in-ili-ʔan	ʔim-pili-ʔan
pres.	ka-ʔum-pilli	ke-pilli-ʔa	ke-ʔippili	ke-pilli-ʔi	ke-ʔi-ppili-ʔi
fut.	ʔum-duyyag	duyyag-en	ʔi-dduyag	duyyag-an	ʔi-dduyag-an
past	d-imm-uyag	d-in-uyag	ʔin-duyag	d-in-uyag-an	ʔin-duyag-an
pres.	ka-ʔum-duyyag	ka-duyyag-a	ke-ʔi-dduyag	ka-duyyag-i	ke-ʔi-dduyag-i

Note first that the nasal of the prefix *ʔum-* is constant, while the nasal in *ʔin-* agrees in point of articulation with the first consonant of the root. We shall return to this nasal assimilation rule later and justify the assumption that the underlying nasal in *ʔin-* is dental.

Secondly, we observe that the medial root consonant is geminated in the non-past of the subject, object, and referent foci, but not in the accessory or beneficial. Since in the latter two cases the initial root consonant is doubled because of the presence of the prefix ?i-, one might imagine that there is an incompatibility between a stem-initial and medial geminate. However, there are other contexts where both kinds of gemination occur, as in the following stative paradigm of bitu 'to put'.

fut.	me-?i-bbittu-?an
past	ne-?i-bitw-an
pres.	ke-?i-bbittu-?an
imper.	?i-bitu

We thus require the following provisional statement of the medial gemination rule: double the medial consonant of a CVCV(C) root in the imperfective except in the accessory and beneficial focus.

We now turn to CVCV(C) roots whose first vowel is e. hepfung 'to break a stick' is inflected as follows.

	subj. f.	obj. f.	access. f.	ref. f.	ben. f.
fut.	?um-hepfung	hepfung-en	?i-hhepfung	hepfung-an	?i-hhepfung-an
past	h-imm-epung	h-im-pung	?in-hepfung	himpung-an	?in-hepfung-an
pres.	ka-?um-hepfung	ke-hepfung-a	ke-?i-hhepfung	ke-hepfung-i	ke-?i-hhepfung-i

Let us discuss the past tense forms first. Like many other Philippine languages, Keley-i has preserved the rule deleting the pepet vowel from the first syllable of a root so long as an initial cluster or a medial three-consonant cluster does not arise. Thus, past tense object and referent forms like h-im-pung and h-im-pung-an from /h-in-epung(-an)/ result from loss of the e followed by the nasal assimilation rule. The rule does not apply in the subject past h-imm-epung because the e is preceded by a cluster nor, for the same reason, in the accessory and beneficial past where the prefix ?in- creates a cluster with the root-initial consonant. Forms like b-in-edbed 'bind' object past show that a following cluster also inhibits the rule. We thus formulate the following rule of syncope, where the + stands for the stem boundary.

e → Ø / V+C\_\_CV

Turning now to the present and future forms, note that a copy of the root-initial consonant is placed after the e. This process is presumably related to the fact that the present and future is formed by reduplication of the first root syllable in other Philippine languages. Compare the following paradigm for sulat 'read' in Tagalog.

fut.	susulat
past	s-um-ulat
pres.	s-um-usulat

We thus assume that the root shape *hehpung* derives from *hehepung* via the syncope rule. This in turn suggests that perhaps the medial gemination process is the reflex of an historical reduplication of the second root syllable. All we need assume is that upon reduplication of a root such as *pili* to *\*pilili*, the medial vowel was reduced to *pepet* (*\*pileli*) - a type of reduplication found in many American Indian languages. Synchronically, however, there is no evidence that the shape *pilli* arises from anything but a rule like medial gemination. We thus have the following three stem modification rules, each restricted to the present and future tenses.

reduplication:	$C_i eCV(C) \rightarrow C_i eC_i eCV(C)$
medial gem.:	$\emptyset \rightarrow C_i / C[i,u,a]C_i \_\_ V(C)$
initial gem.:	$\emptyset \rightarrow C_i / ?i-\_\_ C_i$

Initial gemination may combine with either reduplication or with medial gemination (except in the accessory and beneficial), while reduplication and medial gemination are mutually exclusive. A form such as *?i-hhehpung* is derived as follows.

	/?i-hepung/
reduplication	?i-hehepung
initial gem.	?i-hhehpung
syncope	?i-hhehpung

*CaCV(C)* roots are more complex, as a glance at the inflection of *gabut* 'to cut grass' shows.

	subj. f.	object f.	access. f.
fut.	mang-gebbut	gebbut-en	?i-ggabut
past	nang-gabut	g-in-ebut/g-im-but	?ing-gabut
pres.	ka-mang-gebbut	ka-gebbut-a	ke-?i-ggabut
	ref. f.	ben. f.	
fut.	gebbut-an	?i-ggabut-an	
past	g-in-ebut-an/g-im-but-an	?ing-gabut-an	
pres.	ka-gebbus-i	ke-?i-ggabus-i	

In the imperfective there is a change of the *a* root vowel to *e* whenever the medial gemination rule has applied. This change does not occur before any other consonant clusters - only those that result from medial

gemination (cf. *ʔagtu* 'to head carry'). This alternation lends a modest degree of support to the contention that medial gemination is the reflex of an historical reduplication with reduction to *pepet*, since if *gebbut* is derived from *\*gabebut*, the rule fronting *a* to *e* before *i* and *e* could account for this otherwise strange alternation. Synchronically, however, it is clear that this rule is no longer productive, as it affects only the present morpheme *ka-*. We thus require another special stem modification rule that will switch *a* to *e* when followed by a geminate that has arisen from the medial gemination rule.

Turning now to the past tense forms, note that *a* has changed to *e* in *g-in-ebut* and *g-in-ebut-an*. This is the result of a rule in Keley-1 that raises the first vowel of a *CaCV(C)* root to *e* when that vowel is in the context *VC\_\_CV*. The rule does not apply when the *a* is preceded by a consonant cluster (*ʔing-gabut*, *ʔi-ggabut*) nor when followed by a consonant cluster (*ʔ-in-agtu* 'head carry'). The alternative pronunciations *g-im-but* and *g-im-but-an* occur in more colloquial speech and can be viewed as arising in one of two different ways. First, it is possible that the *e*-syncope rule is being generalised to apply to *a* as well. Alternatively, it is possible that the *a* is first raised to *e* and then deleted in the colloquial style by the *e*-syncope rule. At this point there is no way to decide between these two alternative analyses.

Non-high vowels in the second root syllable are also subject to morphophonemic change, but again only when they are in the context *VC\_\_CV*. In this context *a* and *e* delete. The following imperative and past tense referent focus forms illustrate this aspect of Keley-1 verbal morphophonemics.

<i>gaget</i>	<i>gags-i</i>	'hurry!'
<i>ʔayag</i>	<i>ʔayg-i</i>	'call him!'
<i>dilag</i>	<i>d-in-ilg-an</i>	'lighted'
<i>giked</i>	<i>g-in-ikd-an</i>	'prepared rice field'

This behaviour differs from that displayed by *a* and *e* in the first root syllable in two respects. First, there are no examples in which *a* shifts to *e* and does not delete. Second, there are cases where *a* and *e* fail to delete from the second root syllable even though in the context *VC\_\_CV* (see below); on the other hand, there are no examples where *e* fails to syncope from the first root syllable when in the context *VC\_\_CV*. These facts lead us to believe that there is a separate syncope rule (syncope II) operating in the second root syllable that deletes both *a* and *e*.

[*a, e*] - Ø / *VC\_\_C+V*



Interesting problems arise when we consider referent past roots where both the first and the second vowel are non-high, for both are then in the context for morphophonemic change. In general what we find is that only one of the two syllables is affected - never both. In CeCeC and CeCaC roots, it is always the first root vowel that is deleted.

ʔedep	ʔ-in-dep-an	'extinguished'
degeh	d-ing-geh-an	'was sick'
dengel	d-ing-ngel-an	'heard'
hegep	h-ing-gep-an	'entered'
depap	d-im-pap-an	'wrestled'
ʔekal	ʔ-ing-kal-an	'removed'
getad	g-in-tad-an	'beat gongs'

This may be described by simply ordering the first syncope rule before the second. When the former is applied, a consonant cluster is created, preventing application of the second syncope rule, as shown by the following derivation.

/d-in-epap-an/	
d-in-pap-an	syncope I
inapplicable	syncope II
d-im-pap-an	nasal assimilation

In CaCeC and CaCaC roots two patterns occur. Usually the second root vowel is deleted and the first remains as a. When the second root vowel exceptionally fails to delete, the first a will raise to e but never deletes.

gaget	g-in-agt-an	'hurried'
ʔameh	ʔ-in-amh-an	'was jealous'
takew	s-in-ekew-an	'stole'
gatel	g-in-etel-an	'itched'
gawat	g-in-awt-an	'borrowed'
tapang	s-in-apng-an	'attracted'
hagad	h-in-egad-an	'swept'
gahal	g-in-ehal-an	'scooped'

This behaviour can also be described by simply ordering syncope II before the rule that raises a to e in the first root syllable. Deletion of the second root vowel creates a consonant cluster that prevents raising of the first root vowel. When syncope II exceptionally fails to apply, the a of the first syllable is now in the context VC\_\_CV, and so raising does occur. The following derivations show how these rules work.

/g-in-aget-an/	/g-in-atel-an/	
inapplicable	inapplicable	syncope I
g-in-agt-an	inapplicable	syncope II
inapplicable	g-in-etel-an	raising

Note that if the one-step option of simply deleting a in g-im-but from /g-in-abut/ is selected by generalising syncope I to apply to e and to a (in the colloquial style), then we must order this rule before syncope II when the first root vowel is e, but after syncope II when the first root vowel is a. On the other hand, if the two-step method of raising to e and then reapplication of syncope I is adopted, we again have the same rules applying in different orders: syncope I would occur before raising in the formal style and after raising in the colloquial. Although phonological rules normally apply in the same order for all derivations, Keley-1 seems to be a language in which different orderings occur rather often. The nasal assimilation process presents a similar problem.

Keley-1 contrasts nasal at three points of articulation: labial, dental and velar. When standing before a consonant, n assimilates in point of articulation, while m does not. There are no good examples in which the behaviour of the velar nasal ng can be assessed. It does not unambiguously appear in any prefix or infix and there are no stems of the shape ngeCVC, where, upon deletion of the e, the assimilatory nature of ng could be determined.

The prefix and infix (?)um- illustrates the constant non-alternating nature of m.<sup>3</sup>

teled	?um-tetled	'sting'	?eba	?um-?e?ba	'carry on back'
dengel	?um-dedngel	'hear'	petut	p-um-tut	'dam'
kebed	?um-kekbed	'scratch'	bedad	b-um-dad	'untie'
gelid	?um-geglid	'move'			

The infix -in- assimilates the point of articulation of a following consonant when inserted into a root of the shape CeCVC, since here syncope I operates.

təpen	s-im-pen	'measure'	hemek	h-im-mek	'pity'
kəbet	k-im-bet	'scratch'	tewik	s-in-wik	'prick'
petut	p-in-tut	'dam'	peyuh	p-in-yuh	'bless'
bedad	b-in-dad	'untie'	behat	b-in-hat	'cut rattan'
tekuk	s-ing-kuk	'shout'	de?ek	d-in-?ek	'accuse'
pənel	p-in-nel	'hold'			

Since all Keley-i roots begin with a consonant, prefixes such as the accessory past ?in- and the intransitive nan- never occur before a vowel, so their underlying final consonant can never be directly observed. But since they exhibit exactly the same range of variants as the nasal in the infix -in-, which is unambiguously a dental, they may be safely assumed to also end in a dental nasal morphophonemically.

pehal	?im-pehal	'crack bamboo'			
behat	?im-behat	'cut rattan'	?ala	?in-?ala	'get'
tewik	?in-tewik	'prick'	bayu	nam-bayu	'pound rice'
dengel	?in-dengel	'hear'	degeh	nan-degeh	'hurt'
keleng	?ing-keleng	'sacrifice'	gubat	nang-gubat	'fight'
gitek	?ing-gitek	'cut'	hi?gut	nan-hi?gut	'knot'
heged	?in-heged	'wait'	?awit	nan-?awit	'get'

We thus formulate the following rule of nasal assimilation, ordered after syncope.

$$n \rightarrow \left[ \begin{array}{c} \alpha \text{ point of} \\ \text{articulation} \end{array} \right] / - \left[ \begin{array}{c} + \text{ cons} \\ \alpha \text{ point of} \\ \text{articulation} \end{array} \right]$$

Since the oral and laryngeal glides are [-cons] they will not trigger assimilation and the underlying dental nasal of ?in- and nan- will show up before these sounds.

There is another set of prefixes in Keley-i that invoke a deletion of the root-initial consonant. These prefixes mark the contrastive identification of the agent of the verb and when the subject is pro-nominal, the oblique set is used and the pronoun appears before the verb. Compare ?um-beyyu ?ak ni pagey 'I'll pound some rice' and hi?g-ak mem-eyyu ni pagey 'I'll be the one to pound some rice'. In this set of prefixes, the perfective is marked by nen- and the imperfective by men-: cf. the various forms of gubat 'fight': meng-ubbat, neng-ubat, ke-meng-ubbat. Taking the past tense prefix as the paradigmatic example, note that in the following data it exhibits the same range of alternation as ?in- and nan-.

bayu	nem-eyu	'pound rice'
patey	nem-etey	'kill'
duntuk	nen-untuk	'hit'
gubat	neng-ubat	'fight'
hulat	nen-ulat	'cover'

Assuming that the underlying form of the contrastive identification prefix is nen-, it is clear that nasal assimilation must precede the

rule deleting the root-initial consonant after this type of prefix. But this leads to an ordering paradox. Nasal assimilation precedes root-initial consonant deletion. But the latter rule must precede syncope, as the following forms show.

pedug	nem-dug	'chase'
beka	nem-ka	'dig'
kebet	neng-bet	'pick up'

For it is only by the deletion of the root-initial consonant that the VC CV environment for syncope is created. A form like nem-dug would be derived as follows.

/nen-pedug/	
nem-pedug	nasal assimilation
nem-edug	root-initial consonant deletion
nem-dug	syncope

Note that in the above data the final nasal of the prefix agrees in point of articulation with the deleted root-initial consonant. In the following forms, on the other hand, the nasal agrees with the medial consonant.

teba	nem-ba	'kill pig'	depu	nem-pu	'possessed by spirit'
tepen	nem-pen	'measure'	dengel	neng-ngel	'hear'
temel	nem-mel	'plant sprouts'	hemek	nem-mek	'pity'
teled	nen-led	'sting'	hepaw	nem-paw	'possess'

What accounts for this difference? Clearly it is the fact that the root-initial consonant in these stems is one before which no assimilation of the prefixal nasal occurs and upon the deletion of the e, nasal assimilation applies to assimilate the final n of nen- to the medial consonant of the root. But this implies that nasal assimilation follows syncope, completing the ordering paradox. On this analysis, a form such as rem-ba would be derived as follows.

/nen-teba/	
vacuous	nasal assimilation
nen-eba	root-initial deletion
nen-ba	syncope
nem-ba	nasal assimilation

The Keley-1 ordering paradox crucially depends of course on the identity of the nasal assimilation that occurs in the prefix nen- and that which occurs after the syncope rule. So far they have been exactly identical. There is, however, one factor which calls this identification

into (serious?) question. This is the fact that the nasal in *nen-* assimilates to a root-initial glottal stop, while the other prefixes *ʔin-*, *nan-*, etc. as well as the nasal-consonant clusters resulting from syncope do not.

ʔala	ʔin-ʔala	'get'
ʔawit	nan-ʔawit	'get'
ʔinum	neng-inum	'drink'
ʔeba	neng-ba	'carry on back'
deʔek	nen-ʔek	'accuse'

We thus require a modification in the rule of assimilation for *nen-* so that its final nasal will assimilate to a root-initial glottal stop. But no such assimilation occurs for *ʔin-* and *nan-*. Is this difference enough to justify the conclusion that there are two separate nasal assimilation rules in Keley-i? If so, then the ordering paradox is resolved.

The question of whether two phenomena are part of the same rule or not is of course one of the most difficult issues facing contemporary phonological theory. Until a theoretical clarification of this issue occurs, the Keley-i nasal assimilation problem remains open.

#### N O T E S

1. For example, Keley-i *išu* 'we (incl.)' corresponds to Kalanguya *itayu*. As we shall see, *a* is raised to *e* in certain positions in Keley-i and may then drop out, giving \**ityu*. Also, the frozen form *iyeja* 'it is here' can be related, at least historically, to *deya* 'here' from earlier \**iye-dya*.

2. Gliding of *u* to *w* before a vowel does not take place if a cluster of three consonants would arise. Hence, no gliding occurs when *ʔagtu* is followed by a vowel. Instead, the rule inserting a glottal stop between vowels applies.

3. This morpheme is infixes in subordinate clauses and prefixed in main clauses.

BIBLIOGRAPHY

REID, Lawrence A.

- 1975 'The Pronominal System of Proto-Cordilleran, Philippines'.  
*Working Papers in Linguistics, University of Hawaii* 6/6:  
1-12.

SCHACHTER, Paul

- 1976 'The Subject in Philippine Languages: Topic, Actor,  
Actor-topic, or None of the Above?'. In: Charles N. Li,  
ed. *Subject and Topic*, 491-518. New York: Academic Press.

# SOME FEATURES OF NARRATIVE DISCOURSE IN KADAZAN

HOPE M. HURLBUT

## 0. INTRODUCTION

Some outstanding features of narrative discourse in the Labuk dialect of Kadazan<sup>1</sup> are described in this paper.

Several features of discourse were examined with a view to comparing native text with translated texts. It is hoped that such a comparison will point out areas of special difficulty which would hinder or distort a text resulting in an unnatural translation, and that guidelines can be formed to aid in future translation into the Kadazan language.

During the study of native first person narrative, the author felt subjectively that a translation of the Gospel of Mark seemed to have features that are found exclusively in folktales, so a comparison with features of folktales has been included.

## 1.0. FEATURES OF ACTUAL-EVENT NARRATIVES

The following is based primarily on an autobiographical narrative entitled 'My Marriage'. Reference is made to other narratives where necessary to illustrate certain features of discourse.

## 1.1. OUTLINE OF SEMANTIC CONTENT

The text is a first person report about a marriage ceremony. The story was related in November 1972 by a woman from the village of Kamansi, not far from the mouth of the Labuk River.

Paragraph 1: The narrator introduces herself and where she is from.  
Her husband is introduced and where he is from.  
The husband and friends go upriver to the wife's village.  
He is to go up to the wife's house the next day.

Paragraph 2: They sleep one night.

Many gather.

The custom of carrying the bride and groom throughout the wedding is explained.

Paragraph 3: The next day the ceremony takes place.

The bridal party goes downstream to the husband's house.

Rain comes - all get wet except the bride and groom.

Difficulties of travel in the rain are described - the rain stops on arrival.

Paragraph 4: Many Chinese guests are invited.

Paragraph 5: Description of feasting, drinking and dancing.

Narrator expresses shame at what happened and gives more details.

Paragraph 6: Conclusion.

## 1.2. SURFACE STRUCTURE

Some of the outstanding features of the surface structure of narrative discourse as found in the text 'My Marriage' and corroborated by data from other texts are described and illustrated in the following paragraphs.

### 1.2.1. Groups

"Within a discourse, the content is not presented in an undifferentiated stream, like an inventory or shopping list. Whatever is to be said is grouped, so that related material is together, and the relationship of the groups to each other may be seen." (Callow 1974)

In this study clauses will not be dealt with, but sentences and paragraphs and episodes will be described briefly.

#### 1.2.1.1. Sentences

Sentences in Kadazan are composed of one or more independent clauses, which are linked together.

A new sentence is recognised each time there is a change of participant or there is a change of time or location unless the sentences are conjoined.

The length of sentences varies from one word to as many as 30. The average length is about eight words.

So far 14 sentence types have been set up on the basis of the semantic relationship between the parts of the sentence. On this basis



it was found that the various conjunctions occurred in more than one sentence type, and that each conjunction had a rather wide range of meaning.

Only five of these sentence types occur in our present narrative.

#### Simple Sentence

Tontok diri kumawin oku do ioku nopo nga sodi oku Kamansi. (*Time that marry I that I there I Kamansi.*) 'At the time that I got married, as for me I was at Kamansi.'

#### Non-sequential-juxtaposed

Varo sonsodoppo dii, nokoodoppo dii. (*There-was one-night, slept.*) 'There was one night, they slept.'

#### Merged

Minaja oku diri savoku mongoi sodi valai doisido. (*Followed I that husband-my go there house his.*) 'I accompanied my husband to go to his house.'

#### Co-ordinate

Jadi osusapo dino tumulak vagu poliad di talud, om au nogi antangan do osovutno iri kansang marun-darun. (*So-then difficult that push-off again push-away boat, and not also see for misty that heavy raining.*) 'So then it was difficult to push off the boat again, and one could not see also for that heavy rain made it misty.'

#### Addition

Na mongoi nopo modsu atau pun sitataai atau pun sisobu, na lumopot do nulou om mokibalung do tulun. (*Then going bathe or want-to-defecate or want-to-urinate, then wrap-up with cloth and ask-to-carry-on-shoulders person.*) 'Then as for going to bathe or wanting to defecate, or wanting to urinate, then one wraps up in a cloth and asks someone to carry one.'

#### 1.2.1.2. Paragraphs

Paragraphs are set up on the basis of change of time setting and/or location setting. In the discourse 'My Marriage' all paragraphs are marked initially by a sentence with a temporal margin, and some are closed by a temporal margin also.

#### Paragraph 1

Tontok diri kumawin oku ... (*Time that marry I ...*) 'At the time that I got married ...'

## Paragraph 2

Varo sonsodoppo dii ... (*There-was one-night ...*) '*There was one night ...*'

## Paragraph 3

Jadi korikotpo dii susuab ... (*So-then arrived next-day ...*) '*So then the next day having arrived ...*'

## Paragraph 4

Korikot ikonoï sodi ... (*Arrived we there ...*) '*We having arrived there ...*'

## Paragraph 5

Om kawin nopo gia di silo dagai di pogulu ... (*And marry there ours before ...*) '*And on getting married at our place before ...*'

## Paragraph 6

Om irino dii tontok kumawin ... (*and that-is time marry ...*) '*And that was it at the time of getting married ...*'

1.2.1.3. *Episodes*

The present story 'My Marriage' does not have episodes as opposed to paragraphs, that is each separate episode is contained in one paragraph. There are however episodes in a third person narrative about 'Headman Tinggawi and the Bubua'.

## Episode 1

The story begins with two paragraphs giving the setting and background about the wife before her marriage, and how she became a Muslim.

## Episode 2

There are three paragraphs describing her married life, children, her final illness and death.

## Episode 3

The main episode covers 12 paragraphs telling how her spirit indwelt a bubua (an animal) which came to bother her husband because he caused her to break her Muslim vows. It ends in the death of the bubua.

## Episode 4

The last two paragraphs contain an overlay which tells the whole story again in an abbreviated form, giving the surprising information that the bubua did not die after all.

### 1.2.2. Function of the Groups

Various grammatical structures have been observed to have specific and repeating functions over the length of a text.

#### 1.2.2.1. Sentences can function at both paragraph and discourse levels.

An equational sentence functions to introduce main participants, e.g. *Tontok diri kumawin oku do ioku nopo nga sodi oku Kamansi. (Time that marry I that I there I Kamansi.) 'At the time that I got married, as for me I was at Kamansi.'*

An existential clause has two uses. It introduces subordinate participants, e.g. *Varo iri jaga-jaga-ka. (There-were those care-for-say.) 'There were those who took care, they say.'* It is also used in a time clause to introduce the body of the discourse, e.g. *Varo sonsodoppo dii ... (There-was one-night ...) 'There was one night ...'*

A sentence with a time margin introduces paragraphs. A general time margin begins the introductory paragraph, and in the story 'My Marriage' it also begins the concluding paragraph, e.g. *Tontok diri kumawin oku ... (Time that marry I ...) 'At the time that I got married ...'* A more specific time margin begins other paragraphs, e.g. *Varo sonsodoppo dii ... (There-was one-night ...) 'There was one night ...'*

Certain sentence types occur more frequently in different types of narratives. In the present narrative only five types occur - simple, merged, addition, co-ordinate, and non-sequential-juxtaposed. At the peak the sentences are short, simple ones.

#### 1.2.2.2. Analysis of the surface structure of paragraphs is still incomplete. However semantic functions of the paragraphs in the text 'My Marriage' can be given as follows.

Paragraph 1 functions as the introduction. In it the main participants are introduced and the setting is given for the whole discourse.

Paragraph 2 functions as a build-up one and an embedded procedural discourse is begun.

Paragraph 3 functions as climax/peak and ends the embedded procedural discourse. A further embedded discourse, a travel narrative is included as a digression.

Paragraph 4 functions as post-peak describing one group of wedding guests.

Paragraph 5 functions as a second post-peak describing what guests do at a typical wedding.

Paragraph 6 functions as closure and summarises the discourse.

1.2.2.3. Analysis of the surface structure of episodes is still incomplete. However semantic functions of episodes in the text 'Headman Tinggawi and the Bubua' can be given as follows.

Episode 1 functions as the setting and background of the narrative.

Episode 2 functions as the build-up of the narrative.

Episode 3 functions as the pre-peak and peak of the narrative.

Episode 4 functions as a conclusion.

### 1.2.3. Features of Climax Points

Four features that help to highlight information are outlined here. One or all of these features may occur at the peak, but can also occur at other important points in the narrative.<sup>2</sup>

1.2.3.1. Rhetorical underlining is the repetition of certain words, phrases, clauses and even of whole sentences, sometimes word for word, sometimes paraphrased, which helps to slow down the pace of the story and to bring the repeated item into prominence.

At the peak of the narrative 'My Marriage' the ceremony itself is repeated twice, e.g. Popiurungno dii. Miurungno dii. (*Cause-to-sit-ceremonially-side-by-side. Sit-ceremonially-side-by-side.*) 'They are caused to sit ceremonially side-by-side. They sit ceremonially side-by-side.'

There is repetition of the particle dii which is often used as emphasis, e.g. Jadi korikotpo dii susuvab, popiurungno dii. Miurungno dii. Mitingkuangpo dii kavagu-ka sodi valai diri kusai. (*So-then arrived the-next-day, cause-to-sit-ceremonially-side-by-side. Sit-ceremonially-side-by-side. Return-to-the groom's-house again-they-say there house that male.*) 'So then the next day having arrived, they are caused to sit ceremonially side-by-side. They sit ceremonially side-by-side. They return, it is said, to the house of the man.'

There is also repetition of setting, e.g. Mitingkuangpo dii kavagu-ka sodi valai diri kusai. Minaja oku diri savoku mongoi sodi valai doisido. (*Return-to-the-groom's-house again-they-say there house that male. Followed I that spouse-my go there house his.*) 'They return, it is said, to the house of the man. I accompanied my husband to go to his house.'

Surprise may be indicated also by repetition in the form of a negative-positive paraphrase, e.g. Korikot ikoi sodi valai, na notingkodno diri darun, auso-i diri, iri darun diri, nokorikot ikonoi. (*Arrived we there house, then stopped that rain, nothing that, that*

*rain that, arrived we.) 'We having arrived at the house, then the rain stopped, there was none of that, that rain, when we arrived.'*

#### 1.2.3.2. Heightened Vividness

Both dialogue and rhetorical questions heighten the vividness of narratives.

Dialogue occurs in build-ups to raise the level of excitement. It is in the form of questions and answers or commands which heighten the vividness of the narrative. No dialogue occurs in the narrative 'My Marriage', but the following example is taken from the story of 'The Tandaau' (an animal), e.g. *Jadi "Ai, onu ilo?" kangku. Na kadiolo, "Taau-ka". (So-then "Hey, what that?" say-I. Then say-they, "Don't-know-say".) 'So then "Hey, what's that?" I said. Then they said, "I don't know", they said.'*

Rhetorical questions are frequently embedded in sentences to give heightened vividness to the pre-peak or post-peak information, e.g. *Kuro nong iolo vokon niopos kasangkat do sondiang diri talud do ogumu iri injin sampai tolu. (Why if they others wet all for different those boats for many those engines until three.) 'What about the others, they were all wet for there were different boats for there were many engines up to three.'* *Jadi onu okupo diti do manarino iri savoku om suuon oku diolo manari. (So-then what I-yet this for dances that spouse-my and ordered I by-them dance.) So then what about me for my husband danced and they told me to dance.'*

#### 1.2.3.3. Change of Pace

This is marked by a variation in the size of the constructions and in the amount of connective material.

The sentences at the peak in current event narrative are short, and may be only one or two words at the height of the peak, e.g. *Miurungno dii. (Sit-ceremonially-side-by-side.) 'They sit ceremonially side-by-side.'*

There is a decreased use of sequential connectors between clauses where one would expect them at the peak, e.g. *Jadi korikotpo dii susuvab, popiurungno dii. Miurungno dii. Mitingkuangpo dii kavagu-ka sodi valai diri kusai. (So-then arrived the-next-day, cause-to-sit-ceremonially-side-by-side. Sit-ceremonially-side-by-side. Return-to-the-groom's-house again-they-say there house that male.) 'So then the next day having arrived, they are caused to sit ceremonially side-by-side. They sit ceremonially side-by-side. They return, it is said, to the house of the man.'*

#### 1.2.3.4. *Change of Vantage Point*

This refers to the observer, the one through whose eyes we view the narrative. In the narrative 'My Marriage' the vantage point changes in the denouement from that of the bride and groom to that of the guests. There is a digression in paragraph 5 where the narrator tells of how she and her husband joined in the fun with the guests, but she evaluates her present feelings about her actions, and then returns to the description of the guests and the feast.

#### 1.2.4. Prominence

Prominence refers to any device whatever which gives certain events, participants, or objects more significance than others in the same context (Callow 1974).

1.2.4.1. Paraphrase is saying the same thing in two different ways. There are two ways of paraphrasing, either saying the same thing in different ways in positive statements, or saying it first in a positive statement and then restating it in a negative statement.

In the narrative 'My Marriage' there is an embedded procedural discourse in which sentence 2 paraphrases sentence 1 and sentence 4 paraphrases sentence 3, e.g. *Jadi ioku nopo diri nga nong dagai nopo bansa kumawin, nga lumosok bo. Tolu tadau lumosok, au kovulai-vulai saralom do kalambu. Na mongoi nopo modsu atau pun sitataai atau pun sisobu, nga lumopot do nulou om mokibalung do tulun. Bolungon dino, au buli papanawon si tana-ka, lopoton dino do nulou. (So-then I that then if our race marry, then hide. Three days hide, not can-come-out inside mosquito-net. Then going bathe or want-to-defecate or want-to-urinate, then wrap-up with cloth and ask-to-carry-on-shoulders person. Is-carried-on-shoulders that, not can be-caused-to-walk there earth-say, is-wrapped that with cloth.) 'So then as for me, if it is our race who marry, then one hides. One hides for three days, one is inside a mosquito-net and cannot come out. Then as for going to bathe or wanting to defecate, or wanting to urinate, then one wraps up in a cloth and asks someone to carry one. One is carried on that person's shoulders, one cannot be allowed to walk on the ground they say, one is wrapped up in a cloth.'*

In the same narrative there are examples of positive-negative and negative-positive paraphrase, e.g. *Au diri atantu iri dulung, mikaa-kaa. (Not that certain that prow, go-back-and-forth.) 'The prow was not steady, it went back and forth.'*

1.2.4.2. Repetition of both nuclear and peripheral items of the sentence may occur. It may be word for word repetition or be almost the same.

Repetition of nuclear items occurs at the peak (see 1.2.3.1.).

Repetition of temporal phrases occurs at paragraph breaks, e.g. Jadi sakali korikot ikoi sodi valai, nioposno i tongotulun do kansang i tongodarun. Korikot ikoi sodi valai, nga notingkodno diri darun, auso-i diri, iri darun diri, nokorikot ikonoi.

Korikot ikoi sodi, ogumu jinomput diolo tongoSino rinumikot siri. (*So-then as-soon-as arrived we there house, wet the people for heavy the rain. Arrived we there house, then stopped that rain, nothing that, that rain that, arrived we.*)

Arrived we there, many invited they Chinese arrived there.) 'So then as soon as we arrived at the house, the people were wet for the rain was heavy. We having arrived at the house, then the rain stopped, there was none of that, that rain, when we arrived.

We having arrived there, they invited Chinese to come there.'

Repetition of locatives also occurs, in this case to highlight the difficulties of travel, e.g. ... jadi kosujur ino siri tongorabpa si tosupot. ... kalanggar kobulang di tongokurimbang, i tongotosupot, i rabpa. (... so-then crashed-into that there reeds there undergrowth. ... crashed banged-into banks, the undergrowth, the reeds.) '... so then that crashed into the reeds, into the undergrowth. ... it crashed and banged into the banks, the undergrowth, the reeds.'

1.2.4.3. The particle *dii* occurs frequently in Kadazan texts with several uses one of which is to emphasise. It occurs at paragraph breaks as a type of cataphoric pronoun to show that a new subject is beginning. It also occurs in one type of topicalised phrase. It is found most frequently at the peak. It is also used as an anaphoric pronoun.

At a paragraph break it marks the introduction of a new subject bringing it into prominence, e.g. Varo sonsodoppo dii, nokoodoppo dii. (*There-was one-night, slept.*) 'There was one night, they slept.'

*Dii* is used in some topicalised phrases to increase the emphasis, e.g. Om duduvo ikopoi dii, i pangantin do au niopos. (*And only-two we, the bride-and-groom that not wet.*) 'And it was only the two of us, the bride and groom that were not wet.'

As an anaphoric pronoun the emphasis is not strong, e.g. Iri nogi di kawin dii om nakapanari oku. (*That also that marriage that and danced I.*) 'It was at that wedding also that I danced.'

*Dii* is used frequently at and around the peak (see 1.2.3.1.).

1.2.4.4. The particle *-i* also indicates emphasis. Its domain is very short, as it emphasises only a word or phrase, e.g. *Auso-i diri.* (*Nothing that.*) *'There was none of that.'* *Auso kinotontuan-i.* (*Nothing orderliness.*) *'There was nothing orderly.'* *Kadang-kadang do sumobu-i.* (*Sometimes that urinate.*) *'Sometimes they urinate!'*

1.2.4.5. Unusual use of anaphoric pronouns helps to bring out prominence. The usual anaphoric pronoun used in discourse is *iri* or *diri* which in real life refers to something far away and not visible. In certain instances however one of the other anaphoric pronouns such as *iti* or *diti* *'this'* may be used to highlight some particular person, e.g. *Au oilaan nong tolu-ka iti duvo piniganding iti jungkung.* (*Not know if three-say these two tied-side-by-side these motor-boats.*) *'I do not know if they said it was three or two of these motor-boats that were tied side-by-side.'* *Om korikotpo iti darun nga kansangno.* (*And arrived this rain then heavy.*) *'And this rain arrived, then it was heavy.'* *Jadi onu okupo diti do manarino iri savoku, om suuon oku diolo manari.* (*So-then what I-yet this for dances that spouse-my and ordered I by-them dance.*) *'So then what about me for my husband danced and they told me to dance.'*

In the second example above emphasis is increased also by the word order where the adjective is put in a separate clause. In the third example emphasis is increased by the rhetorical question beginning with *onu* *'what'* and the topicalisation of the pronoun *oku* *'I'*. This type of prominence usually occurs later in the discourse and acts as a device to reiterate the focus of the main theme.

1.2.4.6. Topicalisation refers to bringing a word or phrase forward before the verb to bring it into prominence. Only something that is in focus can be topicalised and so the verb form changes to harmonise with what has been topicalised. It is most frequent early in a discourse as the main participants are introduced in a topicalised equational clause, but it can occur throughout the discourse at important points as anything that is considered important can be topicalised in Kadazan.

Topicalised pronoun, e.g. *Ioku nopo nga sodi oku Kamansi.* (*I there I Kamansi.*) *'As for me I was at Kamansi.'*

Topicalised noun, e.g. *Iri nopo i savoku nga sodi Kalagan.* (*That the spouse-my there Kalagan.*) *'As for my husband, he was at Kalagan.'*

Topicalised verb, e.g. *Na mongoi nopo modsu atau pun sitataai atau pun sisobu, na lumopot do nulou om mokibalung do tulun.* (*Then going*



*bathe or want-to-defecate or want-to-urinate, then wrap-up with cloth and ask-to-carry-on-shoulders person.) 'Then as for going to bathe or wanting to defecate, or wanting to urinate, then one wraps up in a cloth and asks someone to carry one.'*

Topicalised locative from the story of 'Headman Tinggawi and the Bubua', e.g. *Sirino dii mupol-gupol. (There lying-curled-up.) 'There it was lying curled up.'*

Topicalised temporal phrase from the story of 'Headman Tinggawi and the Bubua', e.g. *Sodop-sodopno dii om sumukabaino dii di tukad-ka. (Every-night and climbs steps-say.) 'It was every night then it climbed the steps, they say.'*

#### 1.2.5. Kinds of Information

A narrative discourse potentially has setting, background, collateral information and main events.

1.2.5.1. Setting refers to where, when and under what circumstances actions take place. It includes temporal and locative forms in the discourses studied.

Temporal phrases and clauses refer to chronological setting of events and function to connect the discourse together.

The first sentence in the body of a narrative discourse usually begins with an existential clause containing a temporal word or phrase, e.g. *Varo sonsodoppo dii, nokoodoppo dii. (There-was one-night, slept.) 'There was one night, they slept.'*

Paragraphs usually begin with a time margin. Time margins can also occur at the end of a paragraph giving head-tail linkage between paragraphs. They also can begin embedded paragraphs or discourses, e.g. *Jadi korikotpo dii susuvab, popiurungno dii. (So-then arrived the-next-day, cause-to-sit-ceremonially-side-by-side.) 'So then the next day having arrived, they are caused to sit ceremonially side-by-side.'*

Locative margins normally occur at the end of sentences, rarely in the middle, e.g. *Jadi sumulukno iolo sodi Kamansi. (So-then go-upstream they there Kamansi.) 'So then they went upstream to Kamansi.'*

1.2.5.2. Background information is secondary information such as explanations or comments.

Evaluation is a type of background information, e.g. *Oikuman okuno do auso kinotontuan. (Ashamed I for nothing orderliness.) 'I am ashamed for there was nothing orderly.'*

Dialogue is background information also, but in our present narrative none occurs. The following example is taken from the story 'A Tandaag' (an animal): Jadi "Ai, onu ilo?" kangku. Na kadiolo, "Taa-ka". (So-then "Hey, what that?" say-I. Then say-they, "Don't-know-say".) 'So then "Hey, what's that?" I said. Then they said, "I don't know", they said.'

1.2.5.3. Collateral information, instead of telling what did happen, tells what did not happen (Grimes 1972), e.g. Jadi osusapo dino tumalak vagu poliad di talud, om au nogi antangan do osovutno iri kansang marun-darun. (So-then difficult that push-off again push-away boat, and not also see for misty that heavy raining.) 'So then it was difficult to push off the boat again, and one could not see also for that heavy rain made it misty.'

1.2.5.4. Main events occur in main clauses, as opposed to temporal and locative information which occurs in margins, and the introduction of participants which occurs in topicalised equational clauses or existential clauses. The margins of sentences or connectives provide linkage from one event to the next. Lack of such connections between events indicates a departure from the theme line, a return to the theme-line or a paraphrase of what has gone before. At the peak there are a decreased number of margins and connectors which gives the effect of highlighting the events.

1.2.5.5. Overlay occurs in some biographical narratives in the form of one or more concluding paragraphs to give a summary of the whole story in a shortened form with some details added.

## 2.0. FEATURES OF FOLKTALES

The following is based primarily on a folktale entitled 'The Youth and the Fool'. Reference is made to other folktales where necessary to illustrate certain features of discourse.

### 2.1. OUTLINE OF SEMANTIC CONTENT

The text is a folktale told by the same woman who related the story 'My Marriage' outlined in 1.1. above.

Paragraph 1: The youth and the fool are introduced.

Paragraph 2: The youth goes fishing, his hook gets caught, so he dives in and finds a house full of fish-people.

Paragraph 3: The seven daughters of the fish king and the king are introduced.

Paragraph 4: The king asks for help to remove a hook from one daughter's mouth, so the youth does it.

Paragraph 5: The king gives his seventh daughter to the youth in marriage and they return to land.

Paragraph 6: The youth goes up to the house, the fool asks how to find such a beautiful wife, and the youth tells him.

Paragraph 7: The fool goes and copies all the youth's actions imperfectly and dies as a result.

## 2.2. SURFACE STRUCTURE

Some of the outstanding features of the surface structure of narrative discourse as found in the folktale 'The Youth and the Fool' and corroborated by data from other texts are described and illustrated in the following paragraphs.

### 2.2.1. Groups

Material in a discourse is related together and the relationship of the groups to each other may be seen. Sentences, paragraphs and episodes will each be described briefly.

2.2.1.1. Sentences are composed of one or more independent clauses which are linked together.

A new sentence is recognised each time there is a change of participant or there is a change of time or location unless the sentences are conjoined. A conjoined sentence occurs between two episodes, e.g.

Jadi nokotindal iri tomulok i valai om nokoontong vagu iri kulintagu.  
(*So-then went-up-from-the-water that youth the house and saw again that fool.*) 'So then the youth went up to the house and the fool saw him again.'

Sentence length varies from three words to 45 words with the average about 11 words.

There are seven of the 14 sentence types used in the present narrative. These are simple, non-sequential-juxtaposed, sequential-juxtaposed, antithetical, direct quotation, co-ordinate and addition.

#### Simple Sentence

Jadi varo iso tadau do mongoi iri tomulok mangapon si tuunon. (*So-then there-was one day that went that youth to-fish-with-a-line there*)

*landing-place.) 'So then one day that youth went fishing at the wharf.'*

#### Non-sequential-juxtaposed

Jadi sakalipo dii do naanu disido i tapon om au dii okodong, nakasavit. *(So-then as-soon-as that got by-him the bait and not pull-up, hooked-on-something.) 'So then as soon as he caught something he could not pull it up, it was caught on something.'*

#### Sequential-juxtaposed

Jadi tolongo dii tomulok bibido disido iri kinosovito, tau-tau do nakasavit siri limbungan di valai si saralom di vaig, valai diri raja do sada. *(So-then dives-in youth, unwinds he that place-that-was-caught, what-do-you-know that caught there ridge-pole of house there in water, house that king of fish.) 'So then the youth dived in, he unwound the place that was caught, what do you know it was caught on the ridge-pole of a house in the water, the house of the fish king.'*

#### Antithetical

Jadi ongoyono, ongoyo dii osuo disido iri valai diri, do ogumu do tongosada si saralom, nga okon-i dii sada do tulun-i dii kasangkat-ka iri. *(So-then goes, goes enters he that house that, for many fishes there inside, but not fish for people all those.) 'So then he went, he went into that house, for there were many fish inside, but they were not really fish for they were really people.'*

#### Direct Quote

Jadi "Bulino" kadi tomulok. *(So-then "Can", said-the youth.) 'So then "All right", said the youth.'*

#### Addition

Jadi ingkodomo diti tomulok i mato, om sakali pogingkalat-ka om sirino diolo di bibir di vaig, sii tuunon dii. *(So-then shut-eyes this youth the eyes, and as-soon-as as-soon-as-open-eyes-they-say and there they at edge of water, there landing-place.) 'So then this youth shut his eyes, and as soon as he opened his eyes there they were at the edge of the water, at the wharf.'*

#### Co-ordinate

Jadi nokotindal iri tomulok i valai om nokoontong vagu iri kulintagu. *(So-then went-up-from-the-water that youth the house and saw again that fool.) 'So then that youth went up from the water to the house and that fool saw him again.'*

#### 2.2.1.2. Paragraphs

Paragraphs are set up on the basis of a change of participants or a group of participants.

Paragraph 1 is the youth and the fool.

Paragraph 2 has only the youth.

Paragraph 3 has the inhabitants of the fish king's house.

Paragraph 4 has the youth and fish king.

Paragraph 5 has the youth and king's youngest daughter.

Paragraph 6 has the youth and the fool.

Paragraph 7 has only the fool explicitly mentioned.

Some of the paragraphs are clearly marked at the end with a locative margin or a closure, e.g. *Jadi ino dii. (So-then that.) 'So then that was it.'*

#### 2.2.1.3. Episodes

In folktales there are often two or more clearly marked episodes consisting of one or more paragraphs each. In the story of 'The Youth and the Fool' there are two episodes with a transitional paragraph in between.

Episode 1 consists of 4 paragraphs 2 to 5 telling of the activities of the youth. Paragraph 6 is a transitional paragraph consisting of dialogue and Episode 2 consists of only one paragraph, paragraph 7, telling of the activities of the fool.

#### 2.2.2. Function of the Groups

Various grammatical structures have been observed to have specific and repeating functions over the length of the text.

##### 2.2.2.1. Sentences can function at both paragraph and discourse levels.

An equational sentence functions to introduce secondary participants, e.g. *Iri nopo i tongotama diri tongotongondu diri nga raja do sada. (Those fathers those women those then king of fish.) 'As for the father of those girls, he was the king of the fish.'*

An existential clause is used in two ways. It introduces main participants, e.g. *Kivaro-ka iri sorita pogulu do duvo mikawan-kawan sama-sama kusai, om iso-i tomulok om iso-i kulintagu. (There-was-say that story before of two friends same male, and one youth and one fool.) 'There was a story before about two friends, both men, they say, one a*

*youth and one a fool.*' It also introduces a time margin at the beginning of the body of the discourse, e.g. *Jadi varo iso tadau do mongoi iri tomulok mangapon si tuunon.* (*So-then there-was one day that went that youth to-fish-with-a-line there landing-place.*) '*So then one day that youth went fishing at the wharf.*'

Direct quotation sentences are used in two dialogue paragraphs in this discourse, and as monologue in two other paragraphs. In paragraph 1 the quotation sentences are used in dialogue mainly as repetition to lighten the information load, and also to introduce some new information to carry the theme along. In paragraph 5 the direct quotations are used in the monologue to introduce some new information. In paragraph 6 again there is a dialogue in the form of questions and answers which do not have new information. In paragraph 7 there is one direct quotation sentence as monologue which is imperfectly formed, as it has no closing quotation marker. This increases the feeling of suspense.

Early in the narrative there is a greater variety of sentences. These include all seven types listed in section 2.2.1.1. At the peak and pre-peak there are only simple, addition, and juxtaposed sentences, both sequential and non-sequential.

2.2.2.2. Analysis of the surface structure of paragraphs is still incomplete. However semantic functions of the paragraphs in the text '*The Youth and the Fool*' can be given as follows.

Paragraph 1 functions as the introduction where the main participants are introduced.

Paragraph 2 functions as the first build-up of episode 1.

Paragraph 3 functions as the introduction of the secondary participants.

Paragraph 4 functions as the peak of episode 1.

Paragraph 5 functions as the conclusion of episode 1.

Paragraph 6 functions as a transitional paragraph between episode 1 and episode 2.

Paragraph 7 functions as episode 2 containing build-ups, pre-peak and peak. Its second function is that it carries the pre-peak and peak for the whole discourse.

In this narrative there is no separate conclusion paragraph, but conclusion is provided by the last clause of the last sentence.

### 2.2.2.3. Episodes

Episode 1 functions as the setting and background giving the details necessary for episode 2. It gives the events in which the first main participant and the secondary participants take part.

Episode 2 functions as the pre-peak and peak of the whole narrative. Episode 2 is shorter, more concise and is an imperfect repetition of episode 1 with a different participant and a different ending.

### 2.2.3. Features of Climax Points

Four of the features that highlight information are outlined in this section.

2.2.3.1. Various types of repetition are used at climax points to act as rhetorical underlining.

At the pre-peak there is repetition three times of the word *sakali* 'as-soon-as' in the sentence margin, e.g. Sentence 31: *Jadi sakali irino minongoi diri tomulok dii ...* (*So-then as-soon-as that-is went that youth ...*) 'So then as soon as that youth went ...' Sentence 32: *Jadi sakali "Ingkodomono ino matonu ...* (*So-then as-soon-as "Shut-your-eyes those eyes-your ...*) 'So then as soon as "Shut your eyes ...' Sentence 34: *Sakalipo dii om inurunan isido mangakan ...* (*As-soon-as and swarmed he eat ...*) 'As soon as he did that then he was swarmed and eaten ...'

At three places in the story there is repeated use of the particle *dii*. In paragraph 2 it is used to highlight verbs on the theme-line, e.g. *Tolongo dii, ongoyo dii.* (*Dive-in, go.*) 'He dove in, he went.' It is also used to highlight a surprising piece of information, e.g. *Okon-i dii sada do tulun-i dii.* (*Not really fish for people really.*) 'They were not really fish for they were really people.' In paragraph 4 it is used to emphasise location four times in two adjacent sentences, e.g. *"Ombo dii do siti dii dikovu do kinosovito. Om mongoi oku dii siti, siti dii kinorikotoku," kadi tomulok.* (*"Where for here your place-it-was-caught. And go I here, here place-of-arrival-my," said youth.*) 'Where was it for it was here at your place that it was caught. And I came here, here is the place of my arrival,' said the youth.' It is used extensively in the pre-peak and peak to highlight verbs, negatives and nouns. In the last sentence which is the peak, it is used three times, e.g. *Sakalipo dii om inurunan isido mangakan di sada om napatai dii isido, au dii nokouli.* (*As-soon-as and swarmed he eat by fish and died he, not returned.*) 'As soon as he did that he was swarmed by the fish and eaten, and he died, he did not return.'

At crucial points in the story there is positive-negative repetition, e.g. *Au dii okodong, nakasavit. (Not pull-up, caught-on-something.) 'It could not be pulled up, it was caught on something.'*

There is also repetition of some setting information, e.g. "Ombo dii do siti dii dikovu do kinosovito. Om mongoi oku dii siti, siti dii kinorikotoku," kadi tomulok. (*"Where for here your place-it-was-caught. And go I here, here place-of arrival-my," said youth.*) *"Where was it for it was here at your place that it was caught. And I came here, here is the place of my arrival," said the youth.*

#### 2.2.3.2. Heightened Vividness

Both dialogue and rhetorical questions heighten the vividness of folktales.

Dialogue occurs in build-ups and in the transition paragraph. When it occurs in the pre-peak it is imperfectly formed, and there is no closing quotation marker. This increases the vividness of the narrative, e.g. *Jadi sakali "Ingkodomo ino matonu, nong 'ingkalat' kangku nopo om Ingkalatno." (So-then as-soon-as "Shut-eyes those eyes-your, if 'open-your-eyes' I say, and open-your-eyes.") 'So then "Shut your eyes, if 'Open your eyes' I say, then open your eyes."*

In this narrative, one rhetorical question is used to begin a new paragraph and highlight the beginning of the denouement of episode 1, e.g. *Onupo dii iri janji diri raja do pokowinon disido di tanak disido di pongovian, i rarea diri. (What that promise that king that cause-to-marry he child his the last-one, the maiden that.) 'What about the king's promise for he caused his last child to be married, that maiden.'*

#### 2.2.3.3. Change of Pace is marked by a variation in the size of the constructions, and in the amount of connective material.

In folktales the sentences at the peak tend to be average or extra long, as there is an increased number of events. The sentences contain a serial sequence of clauses with a piling up of verbs telling what the main participant does. The following sentence is the peak of a folktale about 'The Youth and Wild Pigs', e.g. *Nga sakalipo dii do nounus iri gatan om pataamo dii di tomulok mivaja diri kuvo, om pinouli dii di urang tua do bakas i tomulok do nolingasan-i dii iri tulun au kataru, iri bakas au kataru, do mantad diri gatan nokotimporon siri kojuvan disido. (But as-soon-as was-pulled-out that spear and is-thrown-away by youth along-with that what-do-you-call-it, and was-retained by headman of wild-pigs the youth for healed-really that person not well,*



*that wild-pig not well, that from that spear broken-off-inside that body his.) 'But as soon as the youth pulled out that spear and threw it away along with the what-do-you-call-it, and the headman of the wild pigs let him go home because the sick person was healed, that sick pig, which was sick from having a spear broken off in his body.'*

There are fewer connectors at the peak where one would expect them, e.g. *Jadi ingkodomno disido do mato, aupo nakatangar irad "ingkalat-ka", om ingkalatno isido. Sakalipo dii om inurunan isido mangakan di sada, om napatai dii isido, au dii nokouli. (So-then shut-eyes his eyes, not-yet spoken like "open-eyes-say", and open-eyes he. As-soon-as and swarmed he eat by fish, and died he, not returned.) 'So then he shut his eyes, she had not yet spoken like "open your eyes, she said" and he opened his eyes. As soon as he did that, he was swarmed and eaten by the fish, and he died, he did not return.'*

#### 2.2.3.4. Change of Vantage Point

This refers to the observer, the one through whose eyes we view the narrative. In the narrative 'The Youth and the Fool' the first episode is all from the vantage point of the youth. In paragraph 6 however there is a transition, represented by dialogue between the youth and the fool. The youth then drops out of the narrative and episode 2 completes the narrative from the vantage point of the fool.

#### 2.2.4. Prominence

Prominence is a device which gives certain events, participants, or objects more significance than others in the same context.

2.2.4.1. Paraphrase is a feature of sentences in which the same thing is said in two different ways, e.g. *Om mongoi oku siti, siti dii kinorikotoku. (And go I here, here place-of-arrival-my.) 'And I came here, here is the place of my arrival.'*

2.2.4.2. Repetition of both locative and temporal phrases in the sentence margins seems to denote emphasis that is limited to one part of the narrative (see 2.2.3.1.).

2.2.4.3. The particle *dii* occurs as emphasis at various places in the story (see 2.2.3.1. for three uses of *dii*).

It occurs occasionally at paragraph breaks both to end and to begin a paragraph, e.g. sentence 18 (end of paragraph): *Jadi ino dii (So-*

then that.) 'So then that was it.' Sentence 19 (new paragraph):  
 Onupo dii iri janji diri raja ... (What that promise that king ...) 'What about the king's promise ...'

2.2.4.4. The particle -i occurs relatively infrequently and has its domain at phrase or word level. It has the area of meaning of 'really' or 'truly', but is more emphatic and less cumbersome to use than the adverb itself would be, e.g. Okon-i dii sada do tulun-i dii kasangkat-ka iri. (Not-really fish for people-really all-say those.) 'All those were not really fish for they were people, they say.'

2.2.4.5. Another device for giving prominence is the unexpected use of certain anaphoric pronouns. The pronoun *iti* or *diti* 'this' occurs late in the narrative to reiterate the focus of a main participant, e.g. Jadi ingkodomu iti tomulok i mato. (So-then shut-the-eyes this youth the eyes.) 'So then this youth shut his eyes.'

2.2.4.6. Topicalisation is used in equational sentences to introduce new participants or to state some new fact about the participants, e.g. Na ii nopo iri tongosada diri nga varo turu moissusulod-ka tongondou. (Now that those fishes those then there-were seven siblings-they-say female.) 'Now as for those fish then there were seven siblings, females, they say.'

## 2.2.5. Kinds of Information

A narrative discourse potentially has setting, background, collateral and main events.

2.2.5.1. Setting refers to where, when and under what circumstances an action takes place. In the folktales studied it includes temporal and locative forms.

The body of the discourse begins with a temporal word in an existential clause, e.g. Jadi kivaro iso tadau do minongoi iri tomulok mangapon si tunnon. (So-then there-was one day that went that youth to-fish-with-a-line there landing-place.) 'So then one day that youth went fishing at the wharf.'

Other paragraphs may begin with a temporal margin, e.g. Jadi korikot iri tomulok diri ... (So-then arrived that youth that ...) 'So then that youth having arrived ...'

Locative margins normally occur at the end of sentences but can be highlighted in folktales by repetition, and by being moved from the margin of the sentence into a topicalised position, e.g. *Ombo dii do siti dii dikovu do kinosovito. Om mongoi oku dii siti, siti dii kinorikotoku. (Where for here your place-it-was-caught. And go I here, here place-of-arrival-my.) 'Where was it for it was here at your place that it was caught. And I came here, here is the place of my arrival.'*

2.2.5.2. Background is secondary information such as description, explanations or comments, e.g. *Osuka isido diri tanak iri pongovian i taandai. (Likes he that child that last-one the good-one.) 'He liked that last child, the good one.'*

2.2.5.3. Collateral information tells what did not happen, e.g. *Napatai dii isido, au dii nokouli. (Died he, not returned.) 'He died, he did not return home.'*

2.2.5.4. Main events are typically related to each other by sequential connectors. Some temporal and locative margins or clauses occur, and these connect one series of events to the next. Lack of sequential connectors occurs at the peak, before flashbacks and in dialogue.

2.2.5.5. Overlay occurs in some folktales and functions as the conclusion of the narrative.

### 3.0. CONTRASTS BETWEEN THE SURFACE STRUCTURE OF ACTUAL-EVENT NARRATIVES AND FOLKTALES

#### 3.1. GROUPS AND THEIR FUNCTIONS

Sentences, paragraphs and episodes all occur in both types of narrative, but there are differences which are characteristic of each one. There are also differences of function of the groups apparent between an actual-event narrative and a folktale.

3.1.1. Sentences show a number of differences between the two types of narrative.

Sentences tend to be longer in folktales than in autobiographical narrative. In folktales the average sentence is 11 words with some as long as 45 words. In actual-event narrative the average length is eight words, but there are many short sentences of only two or three words,

and the longest one found was 30 words. The longer sentences in folktales probably reflect the fact that the material is familiar and readily understood.

In autobiographical narratives the main participants are introduced in equational clauses, and the secondary participants are introduced in existential clauses, e.g. *Ioku nopo nga sodi oku Kamansi. (I there at Kamansi.) 'As for me I was at Kamansi.'* *Varo iri jaga-jaga-ka. (There-were those taking-care-they-say.) 'There were those who took care they say.'* In folktales this is reversed, the main participants are introduced in existential clauses but secondary participants are introduced in equational clauses, e.g. *Kivaro-ka iri sorita pogulu do duvo mikawan-kawan sama-sama kusai. (There-was-say that story before of two friends, same male.) 'There was a story before, they say, about two friends, both men.'* *Iri nopo i tongotama diri tongotongondu diri nga raja do sada. (Those fathers those women those then king of fish.) 'As for the father of those girls, he was the king of the fish.'*

In actual-event narratives the sentences at the peak are short and simple, but in folktales they are long and complex.

3.1.2. Paragraphs in actual-event narratives usually begin with a temporal margin, but this is rare in folktales. It was much more difficult to recognise paragraph breaks in folktales for this reason. Perhaps because folktales are well-known, temporal sequences are not prominent.

In actual-event narrative some paragraphs function as embedded discourse such as procedural, hortatory, etc. In the folktales studied this did not occur.

Actual-event narratives have one or two final paragraphs that function as a closure tagmeme, but this is optional in folktales.

3.1.3. Episodes in actual-event narratives contain a progression of events over a period of time with major time and location changes marking a change of episode. In folktales, different episodes are recognised when there is a change in major participant. In many folktales the first participant carries out a series of activities to conclusion in episode one, then a second participant repeats the same activities in a little different manner with a different conclusion in episode two.

### 3.2. FEATURES OF CLIMAX POINTS

Three features used to highlight information contrast in the two types of narrative.

### Rhetorical Underlining

Repetition occurs in the periphery of sentences in both actual-event narratives and in folktales, but repetition of words on the event-line at the peak of the narrative only occurs in actual-event narratives.

### Heightened Vividness

Dialogue is used in both types of narrative, but in the pre-peak of folktales the closing quotation markers are dropped which heightens the vividness of the story.

### Change of Pace

The peak of actual-event narratives is marked by extremely short sentences of only two or three words, but in folktales the sentences at the peak are average in length or extremely long.

## 3.3. KINDS OF INFORMATION

One kind of information in actual-event narratives is evaluation which is found in background material. No examples of evaluation have been noted in folktales as yet.

## 4.0. APPLICATION OF DISCOURSE STUDY TO TRANSLATED TEXT

In the light of the study of narrative discourse several verses of the Gospel of Mark were revised. The passage that was revised was Mark 5:1-13.

The first three verses of the chapter were originally translated as follows: <sup>1</sup>Korikot iolo siri sondipau, vonuvo do Gadara. <sup>2</sup>Om tindalno dii i Yisus mantad siri talud, om kivaro iso tulun rinosukan do rogon sinumamung di Yisus do mantad siri kalabangan, <sup>3</sup>do monongijon isido siri kalabangan sabap rinosukan isido do rogon, om au dii kukuoyon di tongotulun mongogos vagu isido, insan kagasanpo do rantai nga au. (<sup>1</sup>Arrived they there opposite-side, country of Gadara. <sup>2</sup>And went-up-from-water Jesus from there boat, and there-was one person possessed by devil met Jesus for from there graveyard, <sup>3</sup>for lives he there graveyard because possessed he by devil, and not able-to-do people tie again him, although tie with chains then not.) <sup>1</sup>They arrived at the other side, the country of Gadara. <sup>2</sup>Jesus went up from the boat, and there was one person possessed by a devil that met Jesus for he was from the graveyard, <sup>3</sup>for he lived in the graveyard because he was possessed by a devil, and the people could not tie him again, even with chains, they could not do it.' In these three verses there are only two sentences. In natural actual-event text the average length of sentences is only

eight words with the longest sentence only having 30 words. The second sentence here is unnaturally long and introduces too much new information. This second sentence has now been divided into three sentences. The information was also rearranged to make it conform more closely to the style of the actual-event narratives examined. <sup>1</sup>Korikot iolo siri sondipau vonuvo do Gadara, om tindalno dii i Yisus mantad siri talud. <sup>2</sup>Kotindal nopo dii isido om kivaro dii iso tulun do sinumamung disido. Iri nopo tulun diri nga rinosukan do rogon, <sup>3</sup>om siri kalabangan ijononjo. Au dii kukuoyon di tongotulun mongogos isido do ogirot isido, do insan kagasanpo do rantai nga au. (<sup>1</sup>Arrived they there other-side, country of Gadara, and go-up-from-water Jesus from there boat. <sup>2</sup>Went-up-from-water he and there-was one person met him. That person that then possessed by devil, <sup>3</sup>and there graveyard living-place-his. Not able-to-do people tie him for strong he, for although tied with chains then not.) <sup>1</sup>They having arrived at the other side, the country of Gadara, when Jesus went up from the boat. <sup>2</sup>When he went up from the water then there was one person met him. As for that person he was possessed by a devil, <sup>3</sup>and there in the graveyard was his living-place. The people were not able to tie him for he was strong, for even tying him with chains was not possible.'

In the revision, the man who is possessed by a devil is identified by an equational clause, instead of the existential clause used in the former version, to conform more closely with the natural texts. The information in the main clause in sentence 1, is repeated as the margin of sentence 2 in order to slow down the information rate. A subordinate clause *do ogirot isido* (*for strong he*) '*for he was strong*' has been used in verse 3 to explain the reason why the man could not be tied. In the original this clause is found in verse 4 only, but in this translation it is used in both verses 3 and 4 to slow down the rate of information and to bring the fact of his strength into prominence. (Compare the prominence given in the similar exclamatory sentence in English: '*He was too strong for anyone to stop him!*' (Mark 5:4 (TEV).)<sup>3</sup>

It was noted in the study of natural texts that very little new information was introduced in dialogue, especially at the beginning of the dialogue which was mainly evaluation and repetition of known information. This was applied to verses 7 and 8. The original translation read: *Om kadi Yisus disido do "Labusno rogon mantad sino tulun-ka", om muvap-luvap i tulun di rinosukan do rogon mitangar di Yisus: "Oou Yisus, tanak do Kinoringan, iri asavat kuasa, onuma koonu-onunu dogo? Mckianu oku dika sii gotuvang do Kinoringan supaya ada ko mangaraat dcgon," kadi tulun rinosukan do rogon. (and said Jesus to-him that:*

*"Get-out devil from there person-he-said", and calling-out the person that possessed by devil speaks to Jesus: "Oh Jesus, child of God, that high power, what have-to-do-with-you me? Beg I you there before God so-that don't you do-bad-to me," said person possessed by devil.) 'And Jesus said to him: "Devil, get out of that person," he said, and the person that was possessed by the devil spoke to Jesus: "Oh Jesus, child of God, the one with great power, what have you to do with me? I beg you before God that you will not harm me", said the person possessed by the devil.'* In the revision Jesus' command has been put into a main clause and the direct quotation no longer carries new information. This also helps to show down the information rate because of the paraphrase. The revision reads as follows: Sinuu di Yisus i rogon lumabus mantad di tulun om kadisido: "Labusno rogon mantad sino tulun-ka". Nga minongoluvap i tulun di rinosukan do rogon do: "Oou Yisus, tanak do Kinoringan, iri asavat kuasa, onuma koonu-onunu dogo? Mokianu oku dika sii gotuvang do Kinoringan supaya ada ko mangaraat dogon," kadi tulun rinosukan do rogon. (*Ordered by Jesus the devil get-out from the person and he-said: "Get-out devil from there person-said." But called-out the person possessed by a devil: "Oh Jesus, child of God, that high power, what have-to-do-with-you me? Beg I you there before God so-that don't you do-bad-to me," said person possessed by devil.) 'And Jesus ordered the devil to get out of the person and he said: "Devil get out of that person," he said. But the person that was possessed by the devil called out: "Oh Jesus, child of God, the one with great power, what have you to do with me? I beg you before God that you will not harm me," said the person possessed by the devil.'*

In the narratives studied it was noted that when new participants were introduced, the number of participants was always stated explicitly if it was known. The number occurred in the same clause as the participants being introduced or in an equational clause immediately following. In the original translation of this story in Mark the pigs were introduced in verse 11 in an equational clause, while the number of pigs was stated at the end of verse 13 in another equational clause:

<sup>11</sup>Om ogumu dii marong iri vogok sonsungal-sungal siri kulud . . .

<sup>13</sup>... Kogugumuo nopo di vogok om duvo noribu motuu ginumu. (<sup>11</sup>*And many very those pigs rooting there hill. . . .* <sup>13</sup>*... The number of pigs then two thousand about many.)* <sup>11</sup>*And many pigs were rooting on the hill. . . .* <sup>13</sup>*... The number of the pigs was about two thousand.'* As noted in section 1.2.2.1. of this paper secondary participants are introduced in an existential clause in an actual-event narrative so in the revision, the pigs are introduced in an existential clause and the

number follows immediately. <sup>11</sup>Kivaro tongovogok sonsungal-sungal sii kulud. Kogumuo nopo dii nga duvo noribu motuu ginumu. (<sup>11</sup>*There-were pigs rooting there hill. The number then two thousand about many.*)  
,<sup>11</sup>*There were pigs rooting on the hill. The number of them was about two thousand.'*

One further change was made in verse 13 when one of the three remaining sentences was divided into two as this seems to be the peak of the episode. Short crisp sentences are found at the peak of actual-event narratives (see under 1.2.3.). However the sentences are still quite long compared to those of the natural texts studied and further study on the language will probably show how these sentences can be made a more natural length.



N O T E S

1. Kadazan is a language that is widely spoken in Sabah, Malaysia. There are many dialects, and the one spoken along the Labuk River has about 12,500 speakers. Many speakers of this dialect are also found in the town of Sandakan, and on the Kinabatangan River. There are variations in the dialect along different parts of the river, but it is mutually intelligible to all speakers. The dialect as it is spoken in Kamansi near the mouth of the river is about 88% cognate with that spoken in Wonod over 100 miles upriver.

This analysis was made in connection with a linguistic workshop sponsored by the Summer Institute of Linguistics in the Philippines. I am indebted to Matius bin Matulang of Wonod, Sabah for editing the narratives and helping with examples. My thanks also go to Elmer Wolfenden and David Thomas of the Summer Institute of Linguistics who made many valuable suggestions during the preparation of this paper.

2. The author received much valuable help from R.E. Longacre's writings when working on this section.

3. Today's English Version of the New Testament, *Good News for Modern Man*. American Bible Society, 1966.

BIBLIOGRAPHY

ASHLEY, Seymour and Lois ASHLEY

- 1971 'Outline of Sentence Types of Tausug'. *Philippine Journal of Linguistics* 2:44-91.

BALLARD, D. Lee, Robert J. CONRAD and Robert E. LONGACRE

- 1971 'The Deep and Surface Grammar of Interclausal Relations'. *Foundations of Language* 7:70-118.

CALLOW, Kathleen

- 1974 *Discourse Considerations in Translating the Word of God*. Grand Rapids: Zondervan.

GRIMES, Joseph E.

- 1972a 'Outlines and Overlays'. *Language* 48:513-24.
- 1972b *The Thread of Discourse*. Ithaca, N.Y.: Department of Modern Languages and Linguistics, Cornell University.

HCLLENBACH, Bruce

- 1975 'Discourse Structure, Interpropositional Relations and Translation'. *Notes on Translation* 56.

LONGACRE, Robert E.

- 1968 *Discourse, Paragraph and Sentence Structure in Selected Philippine Languages*. Vol.1. U.S. Department of Health Education and Welfare. Santa Ana: S.I.L.
- 1976 *An Anatomy of Speech Notions*. Lisse, Belgium: The Peter de Ridder Press.

## THE VERBAL PARTICLE LEU IN THE MAUMERE LANGUAGE

JOAN M. ROSEN

Maumere is a language of central Flores. It was previously studied by Father Arndt (1931). It is the purpose of this paper to present data<sup>1</sup> and a grammatical analysis of part of this language in such a way that the reader can get a feeling for the language as a living thing.

### INTRODUCTION

This paper is based on the general assumption that the verb is the most powerful and influential part of the sentence. Verbal particles are particles which affect the intensity, meaning or scope of the verb. We will try to give a sketch here of the basic functions of the particle *leu*.

In general *leu* seems to function as an intensifying particle in all of its various uses. *Leu* has the following functions in Maumere:

1. It can be used to form the imperative with transitive or intransitive verbs.
2. It can be used to form the comparative (and also can imply a comparative meaning with some verbs - see 5.2.3.).
3. It can intensify adjectives and give them the meaning 'very'.
4. It can be used to emphasise the verb in a general sense.
5. *Leu* with the verb: Particular Meanings

In many but not all cases the connotation of *leu* is bad. The speaker has negative feelings about what is being spoken about. *Leu* seems to have two general functions when added to verbs than can be included in this section:

- 5.1. The speaker or the subject of the sentence disagrees with the action of the verb.  
The reasons for the disagreement of the speaker may be:
  - 5.1.1. The speaker or subject does not like the action.
  - 5.1.2. The subject does not have the right to do what is mentioned.
  - 5.1.3. It is not the proper or expected time for the action to take place.
- 5.2. The action of the verb is done with a specific intention (which usually involves a bad connotation or something considered undesirable by the speaker or subject of the sentence). The following interpretations may appear with this function:
  - 5.2.1. The recipient or goal of the action should be used for another purpose.
  - 5.2.2. The intention of the speaker or subject of the sentence is improper or bad.
  - 5.2.3. To do something instead of someone else (this may not involve a bad intention). This may also involve a comparative idea (see example sentences).

It can be seen that 5.1. and 5.2. above are actually interrelated and not easily separable. Many of the interpretations given under 5.1. and 5.2. may occur with the same usage of a verb and seem to depend on context.

6. Extended uses of *leu*: *Leu* may affect the truth-value of an embedded sentence.
  - 6.1. Embedded sentences may be implied to be false.
  - 6.2. *Leu* may strengthen the truth-value of an embedded sentence.

# 1. *Leu* IN COMMANDS

Commands may be given without *leu* or any other particle, but the feeling is weaker than if the command is given with *leu*.

- (1a) *gide tali ia!*  
*pull rope that*  
*'Pull that rope!'*
- (1b) *gide leu tali ia!*  
*pull part. rope that*  
*'Pull that rope!'*

Sentence (1b) above with *leu* is stronger than sentence (1a) without *leu*.

Sentences (2b), (3b) and (4) below illustrate that leu may be used with transitive and intransitive verbs. Please note that the commands with leu ((2b) and (3b)) are stronger than those without.

- (2a) bano! 'Go!'
- (2b) bano leu!  
go pt.  
'Go already!'
- (3a) lěbe uhe!  
close door  
'Close the door!'
- (3b) lěbe leu uhe!  
close pt. door  
'Close the door!'
- (4) due leu!  
sleep pt.  
'Go to sleep!'

#### An Aside

Another particle [sa<sup>1</sup>.e] (/sae/) may be used in commands in addition to leu. Saie tends to emphasise that the action commanded must be carried out at the moment of the command, whereas leu emphasises that the action of the verb must be done. For example:

- (5a) běli leu!  
give pt.  
'Give it!' OR 'You must give it!'
- (5b) běli saie!  
give pt.  
'Give it right now!'
- (5c) běli leu saie!  
give pt. pt.  
'You must give it right now!'

Sentence (5c) in which both command/imperative particles leu and sae occur, is stronger than both (5a) and (5b).

## 2. Leu IN THE COMPARATIVE CONSTRUCTION

<p>NOUN<sub>1</sub> + (VERB) + ADJ./{ADV. QUANT.} + to?i <u>leu</u> + NOUN<sub>2</sub> '-er than'</p>
---

The general pattern for the comparative construction is given above. It may be read as follows: 'Noun<sub>1</sub> does verb adj-er than Noun<sub>2</sub>' or 'Noun<sub>1</sub> is more adj. than Noun<sub>2</sub>'. Some examples are as follows:

- (6) Goris sEt gaga to.i leu Seve.  
 name paints good looking more than name  
*'Goris paints better than Seve.'*
- (7) meja e.i gēte to.i leu meja iā.  
 table this large more than table that  
*'This table is larger than that table.'*
- (8) a.u gahar to.i ?leu nimu.  
 I tall more than he/she  
*'I am taller than her/him.'*

As is indicated in the pattern above and in example (6), the comparative construction to.i leu may be applied to VERB + ADV. or to VERB + QUANT. Some further examples follow:

- (9) nimu ga gawang to.i leu a.u.  
 he/she eat much more than I  
*'He eats more than me.'*
- (10) rimu pano ropo to.i leu a.u.  
 they walk fast more than I  
*'They walk faster than I.'*
- (11) rimu ?lěbe ěpang to.i leu ami.  
 they play good more than we  
*'They play better than us.'*

The superlative degree does not seem to exist in Maumere. To.i leu can be used to express the superlative degree by contrasting one item or individual with a whole group. For example:

- (12) Averlinus sEt gaga to.i leu mogan sawen.  
 name paint good more than all all  
*'Averlinus paints better than them all.'* OR  
*'Averlinus paints best of all.'*
- (13) nimu gaga to.i leu (ita) mogat sawEt.  
 he/she good more than we all all  
*'He/she is the best of us all.'*

Note: -t = personal ending for the first person plural which is attached to the quantifier.

- (14) au raning to.i leu rimu mogang sawe.  
 you brave more than they all all  
 'You are braver than them all.' OR  
 'You are the bravest of them all.'

Note: -ng = personal ending in Sika dialect, the equivalent of which is -n in other dialects.

### 3. Leu USED TO INTENSIFY ADJECTIVES

Leu has the meaning 'very' when it follows adjectives. As has been mentioned above, it often has negative connotations. This can hold true with adjectives as well as with verbs as the following examples show:

- (15) au plaron leu!  
 you lazy very  
 'You are very lazy.'

Sentence (15) may be used in a complaining sense.

- (16a) au ěpan leu.  
 you good very  
 (16b) au ěpan golo.  
 you good very  
 'You are very good!'

In sentences (16a) and (17a), which follows, leu and the intonation of the sentence are used to convey sarcasm. A sentence with golo 'very' can also convey sarcasm if the intonation of the sentence is appropriate for sarcasm. However, sentences with golo 'very' are more likely to have a positive meaning.

- (17a) nimu bisa leu.  
 he clever very  
 (17b) nimu bisa golo.  
 he clever very  
 'He is very clever.'

In some cases the use of golo and leu may be the same:

- (18a) ami rugi leu.  
 we lose pt.  
 (18b) ami rugi golo.  
 we lose very  
 'We lost very much.'

(18a) is the same as (18b). In this case please notice that the mood of the sentence is negative.

## 4. Leu CAN BE USED TO EMPHASISE THE VERB

In the introduction to this section I mentioned that *leu* in general seems to have an intensifying function. In commands the mood of the command with *leu* seems to be stronger than without *leu*. In the comparative *leu* helps to convey the idea that one thing is stronger in some quality than another, and *leu* conveys the idea of 'very' with adjectives. It should be clear that in each of these cases the quality of *leu* is strong and it functions to make the predicate stronger in degree. In this section I will give examples in which *leu* is used to emphasise the verb, that is to make the meaning or feeling of the verb stronger in a general sense. In the following section (5) I will illustrate the particular forms and meanings this general emphasising function can take.

(19a) *ina piara ami da.a gëte.*  
*mother take care of us until big*

(19b) *ina piara leu ami da.a gëte.*  
*mother take care of pt. us until big*  
*'Mother takes care of us until we are big.'*

In sentence (19a) the feeling is normal, i.e. it is mother's duty to take care of us. In sentence (19b) the verb is stressed and the feeling is that mother takes care of us intentionally and that she really looks after us well. (Note: the idea of intentionally may make it possible to also classify this and other sentences under 5.2.)

(20a) *a.u ata surat odi, gëru au odo a.u pano.*  
*I read letter first then you order me go*  
*'I will read the letter first, then you (can) order me to go.'*

(20b) *a.u ata ?leu surat odi, gëru au odo a.u pano.*  
*I read pt. letter first then you order I go*

The feeling behind sentence (20b) is 'I am really going to read this letter before I go.'

(21a) *a.u hama pëti.*  
*I step on box*  
*'I step on the box.'*

(21b) *a.u hama ?leu pëti.*  
*I step on pt. box*  
*'I really step on the box.'*

In (21b) *leu* tends to intensify the verb, whereas (21a) is just a normal statement.



- (22) a.u u.a ?leu ro.o matean, nimu wi tukang ga.  
 I work pt. almost die he rel.pron. one who eats  
*'I work myself almost to death, and he does the eating.'*

In sentence (22) *leu* conveys the idea that the speaker does an extraordinary amount of work.

- (23a) a.u tama ?rehi, loning au lěbe ba.a uhe.  
 I enter not able because you shut already door  
*'I cannot get in, because you already shut the door.'*

- (23b) a.u tama ?rehi, loning au lěbe leu ba.a uhe.  
 I enter not able because you shut pt. past door

The difference between (23a) and (23b) is that in (23b) the verb is stressed.

- (24) ami odo leu rimu lulu rang, iana rimu lopa  
 we order pt. them first motion pt. so that they not  
 lat tama sěkola  
 late enter school  
*'We order them to go first, so that they won't be late for school.'*

Note: lulu = 'to go first, to be first'

lopa = also used in the negative imperative 'don't'.

In (24) *leu* is used to stress the verb.

- (25a) ita ga.It leu ba.a dadi ita rehi wuleng  
 we want pt. already therefore we not able oppose  
 waler walong  
 any more/again  
*'We already want to, therefore we can't oppose any more.'*

- (25b) ita ga.It ba.a dadi ita rehi wuleng waler  
 we want already therefore we not able oppose  
 walong  
 again

The difference between (25a) and (25b) is that (25a), the form with *leu*, can be read as 'we really want to ...'.

- (26) au buhe leu, rimu ganu la.eng pano, di la.eng.  
 you lie pt. they as if not yet go indeed not yet  
*'You lied, they haven't gone yet, indeed they haven't.'*

The implication of sentence (26) is that the person addressed said that they had already gone. *Leu* is used to stress the verb 'lie'. The speaker in this sentence is slightly angry.

Sometimes *leu* may be used to imply that extra effort was put forth in the execution of the action of the verb as in sentence (27) below:

- (27) nimu mēnang leu hadia.  
       he win pt. gift/prize  
       *'He won the prize.'*

Other people took part in the contest, but only he won the prize because he really tried hard and he succeeded.

## 5. Leu WITH THE VERB: PARTICULAR MEANINGS

It should be noted that in this section many of the interpretations may overlap or be applied simultaneously to the same sentence, and that a great deal of what determines how a particular verb or sentence is interpreted depends on the native speaker's interpretation of a non-linguistic context or situation.

5.1. The speaker or the subject of the sentence disagrees with the action of the verb.

5.1.1. The speaker or subject disagrees, because he or she does not like the action.

- (28a) rimu no:na pare e uma amIn.  
       they plant rice in garden our  
       *'They plant rice in our garden.'*
- (28b) rimu no:na leu pare e uma amIn.  
       they plant pt. rice in garden our  
       *'They plant rice in our garden.'*

Sentence (28a) is a simple statement of fact, whereas in (28b) the action of the subject *'they'* is in opposition with the desires of the speaker. The speaker disapproves and feels they should not plant in her or his garden. (There is also a comparative idea in this sentence, since (28b) can also be interpreted to mean their work, the planting of rice, precedes the speaker's.)

- (29) nimu pakEt leu, labu ha buluk.  
       he wears pt. clothes one short  
       *'He is wearing clothes that are short.'*

Note: ha = rel.pron.

There are several possible interpretations for this sentence: a) *'he is wearing clothes that are not appropriate with what I asked him to wear'*, or b) *'the clothes he is wearing do not belong to him, i.e. he does not have the right to wear them.'*

- (30a) *nimu liʔi hoang a.ung.*  
*he choose money my*
- (30b) *nimu liʔi leu hoang a.ung.*  
*he choose pt. money my*  
*'He chooses my money.'*

Sentence (30a) is merely a statement of fact, whereas in sentence (30b) the feeling that the speaker does not agree with his action is conveyed. The verb is stressed in (30b) and the idea is conveyed that 'after he selected my money he did not want to return it to me.'

- (31a) *nimu hěna aja a.ung.*  
*he fry corn my*  
*'He/she fries my corn.'*
- (31b) *nimu hěna leu aja a.ung.*  
*he fry pt. corn my*  
*'He/she fries my corn.'*
- (31c) *nimu hěna běli aja a.ung.*  
*he fry for corn my*  
*'He/she fries my corn for me.'*

Note: *běli* = benefactive

(31a) is a normal statement. In (31b) the speaker does not agree with or like the subject's action and the idea is conveyed that the speaker did not order him or her to fry his corn. In (31c) the speaker is happy that the subject did something for him/her. The benefactive word *běli* is used in contrast with *leu*.

#### 5.1.2. The subject does not have the right to do what is mentioned.

- (32) *rimu odo a.u ala ʔleu (X).*  
*they order me take pt.*  
*'They order me to take (something).'*

Note: (X) = a deleted object. Maumere seems to frequently omit the object if it is understood instead of using a pronoun in many sentences.

The use of *ʔleu* here conveys the idea that they are ordering me to do something which I have no right to do.

- (33a) *rimu těri olang a.un.*  
*they sit place my*
- (33b) *rimu těri leu olang a.un.*  
*they sit pt. place my*  
*'They sit in my place'.*

In the sentence with *leu*, (33b) above, the idea is conveyed that they do not have the right to sit in my seat.

- (34) *me kěsik ha tia loka leu ara.*  
*child small one that throw out rice*  
*'That small child threw out the cooked rice.'*

The speaker does not agree with the child's action and feels the child should not have thrown out the rice in sentence (34).

- (35a) *au go.a hai ara nimun?*  
*you eat who rice poss.pron.3rd person*  
 (35b) *au go.a leu hai ara nimun?*  
*you eat pt. who rice hers/his*  
*'Whose rice did you eat?'*

(35a) is just a simple question. In (35b) the feeling is conveyed that the speaker disapproves and that the person addressed was wrong to eat the rice since she/he did not have the right to eat it.

- (36) *nimu gata leu surat tia.*  
*he read pt. letter that*  
*'She/he read that letter.'*

She/he did not have the right to read the letter, since it was not for her/him in sentence (36).

- (37a) *nimu nala pInsil a.un.*  
*he take pencil my*  
 (37b) *nimu nala leu pInsil a.un.*  
*he take pt. pencil my*  
*'She/he takes my pencil.'*

In sentence (37b) the speaker feels annoyed that she/he took his pencil, and does not agree with the subject's action. The subject did not have the right to take the pencil.

- (38) *nimu ga?i na?o, ko ele dadi loning nimu gita*  
*she/he want steal but not happen because she/he see*  
*leu a.u.*  
*pt. me*  
*'She/he wanted to steal (something), but it didn't happen because she/he saw me.'*

The implication of sentence (38) is that the subject did not really want to see me.

- (39) *a.u něni ?leu buku nimung ha, loning nimu norang*  
*I ask for pt. book his one because he be/exist*  
*buku ?lele rua.*  
*book mea.wd. two*  
*'I asked for one of his books, because he had two.'*

The feeling behind sentence (39) is that the person did not want to loan me a book or it may have been improper for me to ask for the book, but he/she loaned me the book anyway.

5.1.3. The speaker disagrees with the action of the verb, because it is not the proper or expected time for the action to take place.

(40a) nimu nēni a.u pano.  
he ask me go

(40b) nimu nēni leu a.u pano.  
he ask pt. me go  
'He asks me to go.'

In (40b) the idea is conveyed that it is not time for me to go.

(41) ami nēni leu nimu ele sēkola ena tei.  
we ask pt. him not school earlier this  
'We ask for him not to go to school today.'

The feeling behind sentence (41) is that it is not the proper time for him/her to miss class and that the subject of the embedded sentence really wants to go to school so that our request is against his wishes. (An element of the preceding section 5.1.1. can also be found in this sentence.)

(42) moat, le.e dopo leu a.u di, eong ha a.u  
old man not want call pt. me emphatic pt. not one I  
di naha moga.  
pt. must also  
'Old man, you do not want to call me, but if you called me, I would surely follow.'

The idea behind this sentence is that it is not the proper time for the speaker to go with the person addressed, an older man, but the speaker is willing to go. An interpretation of the sentence could be: 'in fact I will still be sleeping, but if you want to meet me, please wake me up.' The tone of the sentence is not one of anger or displeasure as in most of the examples given in the preceding section (5.1.1.).

(43a) rimu odo a.u kantar.  
they order I sing

(43b) rimu odo a.u kantar leu.  
they order I sing pt.  
'They order me to sing.'

(43a) is a simple statement. In (43b) the idea is that it is not the proper time for me to sing, but they order me to do so. The speaker

is not quite prepared to sing, but is willing to do it. It is also not the speaker's turn to sing. They are asking the speaker to do something out of order.

## 5.2. The action of the verb is done with a specific intention

Some of the examples in this section may convey the idea of something done intentionally in contrast to something which is not done intentionally, such as in the following sentence (44b):

(44a) nimu dĕna a.u bĕler.  
he make me tired

(44b) nimu dĕna leu a.u bĕler.  
he make pt. me tired  
'He makes/made me tired.'

In (44a) the action may be done intentionally or just by chance. However, in (44b) the implication of the sentence is that the action of the verb was done on purpose: *'He ordered me to do something so that I would become tired.'*

However, in most of the examples of leu which I have found the basic meaning of a verb followed by leu in contrast with a verb which is not followed by a verbal particle does not seem to be that the action is done intentionally in contrast with action which is not marked as to intentionality, but the leu seems rather to convey the idea that the action of the verb is done with a particular purpose in mind. For example:

(45a) a.u mata jarang.  
I tie up horse

(45b) a.u mata ?leu jarang.  
I tie up pt. horse  
'I tie up the horse.'

Sentence (45a) is merely a statement of fact, whereas in (45b) the idea is conveyed that the speaker or subject of the action had a particular reason for tying up the horse. Other examples are:

(46a) guru kela ba.a e surat dĕna ele hulir.  
teacher write already in book so that not forget

(46b) guru kela leu ba.a e surat iana dĕna lopa  
teacher write pt. already in book so that make not  
hulir.  
forget

*'The teacher writes in the book so as not to forget something.'*

Note: iana děna ... lopa is better than dena ... ele. However both are acceptable.

(46a) is a simple statement of fact. In (46b) the idea conveyed is that the teacher had a special reason to do the verb, perhaps he or she had never written in that book before.

(47a) nimu    dēna    a.u    rugi.  
          he    make   me    lose

(47b) nimu    dēna    leu    a.u    rugi.  
          he    make   pt.   me    lose

*'He makes/made me lose.'*

In (47a) the idea conveyed is not that he intended to make me lose, but because he took part in the game or contest, I lost. In (47b) the idea conveyed is that he intentionally tried and succeeded in making me lose.

(48) nimu    dēri    leu    olang    a.un.  
       she/he sit   pt.   place   my

*'She sits in my place.'*

The idea in the sentence immediately above is that she sat in my place on purpose. It should be noted that leu meaning 'done with a particular purpose in mind' can occur with verbs that cannot be interpreted as accidental if occurring alone. For example, in (45a) above we would not normally be able to interpret mata 'to tie up' as an act which is done involuntarily. Therefore, I infer that the use of leu in this section is not to contrast voluntary vs. involuntary acts, nor to contrast acts done on purpose vs. those done accidentally, but leu is often used to emphasise the verb and to imply that the action of the verb was done for a specific reason on purpose.

Several different interpretations may occur with leu when it is used to indicate that the action of the verb was/is done with a particular intent in mind:

5.2.1. The recipient or goal of the action should be used for another purpose.

(49a) ami    tea    pare    tia.  
       we    sell   rice   that

(49b) ami    tea    leu    pare    tia.  
       we    sell   pt.   rice   that

*'We sold that rice.'*

The interpretation most frequently given in sentence (49b) is that the rice was intended for another purpose, probably for the family to eat, but for some particular reason of our own we sold it. This sentence may, but does not necessarily, convey the idea that we did not have the right to sell the rice. If we contrast (49b) with (49a), we will see that as in most of the other examples given so far, the verb form without *leu* is used in simple statements of fact and does not have special implications. Much of the emotional force of sentences such as these seems to rest with the verbal particle.

Another example of the same type is:

- (50) *rimu howe leu wair ei gēlas wale une.*  
*they pour pt. water in glass inside*  
*'They pour water into the glass.'*

The implication in sentence (50) is that the water is intended to be used for another purpose.

#### 5.2.2. The intention of the speaker or subject of the sentence is bad.

In many if not all cases where *leu* occurs, one gets a negative feeling from the sentence. Often the speaker seems unhappy about the events discussed or disapproves of them. In many sentences which convey the idea of something done with a particular intention, the intentions of the actor(s) of the sentence are interpreted to be improper, bad, or not beneficial to the speaker. Some examples of this follow:

- (51) *a.u tuUng ?leu nimu.*  
*I escort pt. him/her*  
*'I escort him.'*

The feeling behind sentence (51) is that the speaker had a special reason for escorting him or her somewhere, i.e. that the presence of the person was not really desired. Probably the person was intentionally escorted away from a particular place, although he or she really wanted to stay.

- (52) *nimu dopo leu a.u.*  
*she/he call pt. me*  
*'He calls me.'*

The understanding behind sentence (52) is similar to that of (51). He calls me for a special reason, i.e. my presence is not desired in a particular place. He calls me so that I will leave the place where I am not wanted, although I really want to stay there.



- (53a) rimu rēti surat a.un.  
 they carry off book my
- (53b) rimu rēti leu surat a.un.  
 they carry off pt. book my  
 'They carry off my book.'

In (53a) the implication is that they took my book accidentally, whereas in (53b) they took my book on purpose and their reason for taking the book was not good for the speaker. They had some bad intention toward the speaker when they took the book.

As mentioned in 5.2. the intent of the speaker or subject in doing the action of the verb may not always be bad. In some cases the subject may have a good reason for her/his action, as in sentences (54) and (55). These cases, however, seem to be in the minority.

- (54) rimu odo a.u to leu.  
 they order me laugh pt.  
 'They order me to laugh.'

In sentence (54) they order me to laugh, because they see that I am sad and they want to lift my spirits.

- (55) ami odo leu rimu lulu rang, iana lopa lat  
 we order pt. they go first motion pt. so that not late  
 tama sēkola.  
 enter school  
 'We order them to go first, so that they will not be late to school.'

In sentence (55) the intent of the speaker is also benefactive and not malfactive toward the objects of the action.

### 5.2.3. To do something instead of someone else.

The sentences in this section often seem to contain an implication of comparison, and the function of leu here may be related to the comparative function mentioned in section 2. (Please note that in the following examples (56), (57), (58) leu follows the embedded verb.)

- (56) rimu odo a.u pano ?leu.  
 they order me go pt.  
 'They order me to go.'

The interpretation of (56) can be as follows: 'Someone else has a chance to go, but they order me to go instead. I'm really not prepared to go, but they regard me as better suited for doing something (the act to be done after the person goes) than someone else who would ordinarily be chosen to go before me.'

- (57) rimu odo a.u dor ?leu.  
 they order me answer pt.  
 'They order me to answer.'

The interpretation of (57) is as follows: this may take place in a classroom where the speaker's classmates say, 'It's better if you answer the question asked by the teacher instead of us.' The special intent (see section 5.1.2.) behind their saying this to the speaker may be 'so that the teacher will stop asking questions.' We can see here as in almost all the examples of leu that have been given that many of the interpretations of leu overlap.

- (58) rimu odo a.u ?lɛbe ?leu.  
 they order me play pt.  
 'They order me to play in someone else's place.'

As in examples (56) and (57) above, the leu here also conveys the idea that 'they', the subject of the main verb, think that the subject of the embedded verb, 'I', can do the action of playing better than someone else.

At the beginning of this section I mentioned that these examples often seem to convey the idea that several nouns or verbs are being compared. This aspect of leu seems to be more clearly illustrated by the following:

- (59a) a.u ulu rimu.  
 I before them  
 'I leave before them.'
- (59b) a.u ulu ?leu rimu.  
 I before pt. them  
 'I go before them.'
- (59c) a.u ulu rimu kɛsɪk ha.  
 I before them little one  
 'I go before them a little.'

Note: (59c) is better than (59a).

The important thing to note here is that sentence (59b) implies that 'we started out together, but I arrive first.' In (59b) where leu follows the verb there seems to be the idea of comparison. However, in (59c) the implication is 'we didn't necessarily start out together.' (59c) seems more to be a simple statement that the speaker arrived shortly before some other people, without a strong implication of comparison.

## 6. EXTENDED USES OF leu: CASES IN WHICH leu MAY AFFECT THE TRUTH-VALUE OF AN EMBEDDED SENTENCE

### 6.1. Embedded sentences may be implied to be false.

With some verbs of speaking and reporting, and with the verb *pěrsaiya* 'believe', leu may be used to imply that an embedded sentence is false. For example:

- (60) *rimu wěta leu a.u pano.*  
*they say pt. I go*  
*'They said I left.'*

The implication of sentence (60) is that the speaker is still in the place from which he/she is said to have left, and that the embedded sentence is false. A further implication is that they lied: *'They said I had left to some people who came to look for me, because they didn't want me to meet my guests.'* The second implication would fall under the heading of a special intent on the part of the subjects which was not beneficial to the speaker (see 5.2.2.).

Please notice in sentence (60) that leu follows the main verb of the sentence. Some further examples of the same type are:

- (61) *rimu tutur leu, wěta a.u paro.*  
*they say/speak pt. say I go*  
*'They said that I had gone.'*

The implications of (61) are the same as those of (60). The primary implication is that they lied about the speaker's having already left.

- (62) *rimu kiring leu ami, wěta rimu gahu.*  
*they report pt. us say they hot*  
*'They reported to us that they were sick.'*

(Please notice that again here as in (60) and (61) leu follows the main verb.) In sentence (62) the use of leu is to cast doubt on the truth value of the embedded sentence. The speaker is not sure that they were really ill. It could be a false report.

- (63) *rimu kiring leu ami, wěta rimu pano.*  
*they report pt. us say they go*  
*'They reported to us that they left.'*

In sentence (63) the implication is again that it could be a false report. In fact they may not have gone.

- (64) ita pǝrsaiya leu tutur nimung, ra.Ik eong odi ita  
 we believe pt. speech his if not later we  
 tǝri da.a waung.  
 sit until evening

*'We just say we believe him, because if we do not do that we will have to sit here until evening.'*

In sentence (64) the speaker(s) are really unwilling to believe what he says, but rather than stay longer and to make things easier for themselves they just say they do. This example is related to 5.1.1. in which the speaker does not like the action of the verb. A further implication of sentence (64) is that whatever 'he', the person referred to, said is not true, or the speaker does not believe it to be true.

- (65) a.u tutur ?leu ?wǝta nimu ele norang.  
 I speak pt. say she not be/exist

*'I said she wasn't here.'*

The implication of sentence (65) is that the speaker is lying and the person referred to really is present.

- (66) a.u ?lapor ?leu, le guru, wǝta nimu gahu.  
 I report pt. to teacher say he hot

*'I reported to the teacher that he was sick.'*

The implication of (66) as in (62) is that the speaker is lying and he really is not sick.

## 6.2. Cases in which leu strengthens the truth-value of an embedded sentence.

This function of leu seems to be related to 4, the intensifying function of leu in which it emphasises the verb. Some examples are as follows:

- (67) rimu pǝrsaiya ami na.o leu hoang.  
 they believe we steal pt. money

*'They believe that we are the ones who stole money.'*

In the sentence above the function of leu is to make the verb 'steal' stronger, and to increase the conviction that we (the speakers) really stole the money. The use of leu increases the possibility that the speakers really stole the money. Please note that in this example the leu follows the embedded verb.

Other examples with a semantic effect similar to that of (67) are:

- (68) a.u taser ?leu odi nimu naha plari nan.  
 I guess pt. later he must run pt.

*'I guess later he will have to run away.'*

In (68) the function of leu is to increase the truth-value of the embedded sentence and to create the feeling that the embedded sentence is more likely to be true. There is a greater possibility that he will really have to run away.

- (69)   ami   mEteng   odi    toma    leu   walong   ngawung   tia.  
          we    hope    later   receive   pt.   back    thing    that  
           *'We hope we will get that thing back later.'*

The function of leu in sentence (69) above is to emphasise the verb *toma* 'receive' and to create the feeling that the speakers will really get their property back.

#### SUMMARY

In this paper I have tried to illustrate the functions of the verbal particle leu in Maumere, a language of central Flores. It can be used: 1) in commands, 2) in the comparative construction, 3) to intensify adjectives, 4) to emphasise the verb, and 5) to give special meanings to the verb. In general the character of this particle is strong and often negative.

# NOTES

1. Validity of the data: each sentence in this paper has been found grammatical by at least two informants from different dialect areas, and in general there seems to be a consensus of opinion on the various functions of *leu*.

2. Explanation of the transcription used in this paper:

- a, e, i, o, u : the line under the vowel indicates that the vowel is murmured. Murmured vowels are phonemically distinct from non-murmured vowels in Maumere.
- ?l, ?r, ?w : the consonants which are preceded by a question mark are laryngealised, and are phonemically distinct from their non-laryngealised counterparts.
- /i/, /e/ : these are slightly lower than their English counterparts.
- [E] : this is an allophone of /e/, which is slightly higher than the English /ε/.
- [I] : this is an allophone of /i/.

3. Syllable boundaries occur between two contiguous vowels in the same word. I have only indicated this in a few cases to avoid confusion, for example *ba.a* 'already' consists of two syllables as does *ga.it* 'we want'. *leu* [le.u] is also a two syllable word.

Occasionally a glottal stop may be heard between two vowels, but not always. The question mark between two vowels is used to indicate a glottal stop.

If the second of two contiguous vowels is murmured, the transition between the two vowels will be heard as a glide. For example the particle [*sa*<sup>i</sup>.e] mentioned above is phonemically /*sa.e*/.

BIBLIOGRAPHY

ARNDT, Paul

1931      *Grammatik der Sika-Sprache*. Ende, Flores.





# THE USE OF DEICTICS TO IDENTIFY PARTICIPANTS IN BAHASA INDONESIA

DONALD F. BARR

## 0. INTRODUCTION

In any discourse one major consideration is how participants are brought onto the stage of the discourse and subsequently tracked through the text.<sup>1</sup> Distinct patterns and strategies exist in every language to accomplish this process. Bahasa Indonesia has a textual use of deictic words, *ini* 'this', and *itu* 'that', among its repertoire of tracking devices. The purpose of this paper is to show how such deictic words function in Bahasa Indonesia to help identify participants in narrative discourse. Further, it is shown that no analysis of deictics in Bahasa Indonesia can be considered complete if it fails to take into account certain aspects of the total discourse in which deictics occur.

## 1. THE BASIC PATTERN

In Bahasa Indonesia there is a rather standard pattern of introduction and subsequent tracking of participants in narrative discourse. Simply put this is as follows:

- 1) A participant is introduced by a formula: *se-orang* 'one-person', + NOUN PHRASE, often further described in clauses introduced by *yang*.<sup>2</sup>
- 2) The next reference is usually by NOUN PHRASE + *ini*.
- 3) Further reference is by NOUN PHRASE + *itu*.
- 4) Subsequent reference is by means of pronouns, NOUN PHRASE + *itu*, affixing of *-nya* to the participant, use of *tadi* or *tersebut*.<sup>3</sup>

This basic pattern can be seen illustrated in a short story by Mohammed Kassin, 'Cara Chicago', one of several in his collection of stories entitled *Teman Duduk*. In 'Cara Chicago' the main participant,

a woman, who is never given a name, is introduced and further identified as follows:

- (3) Se-buah auto tiba dimuka station dan se-orang  
 one (numeral car arrive in front of station and one-person  
 class. for car (num.class.  
 for person)

perempuan yang berbadan kecil molek berdandan bagus dan  
*female who has body small pretty dressed well and*

beperhiasan cukup, keluar dari auto itu, diiringkan oleh  
*wearing jewelry enough got out from car that accompanied by*

se-orang anak perempuan umur kira-kira delapan tahun.  
*one-person child female age about eight years*

'An automobile arrived in front of the station and a woman who was small of build and pretty, well dressed and wearing a full set of jewelry, got out of the car accompanied by a girl aged about eight.'

- (4) Melihat perempuan yang baru datang ini, kedua laki-laki itu  
*seeing woman who just came this both men that*  
 bergamit-gamitan dan waktu perempuan itu pergi kemuka  
*nudged each other and when woman that go in front of*  
 loket akan membeli karcis, se-orang diantaranya  
*ticket window in order to buy ticket one-person in-among-them*  
 datang menghampiri dia.  
*came approached her*  
 'Seeing the woman who had just come, the two men nudged each other, and when the woman went to the ticket window to buy a ticket, one of them approached her.'

In sentence (3) the woman is rather fully introduced, beginning with the formula *se-orang perempuan* 'one-person female' followed by a rather long descriptive clause introduced by *yang*. In (4) she is identified as *perempuan yang baru datang ini* 'woman who just come this', at the beginning of the sentence and by the end of the sentence she is marked as a well established participant as *perempuan itu* 'woman that', and finally *dia* 'her'. That she is well established is shown by the shorter reference, a single noun rather than a fuller phrase, plus being marked by *itu* 'that', and further in the pronominal reference.

Thus we see this basic pattern of participant introduction and identification, a three step pattern of establishing participant identity. We see the rather set formula of introduction. Subsequent reference is then by means of marking the noun phrase identifying the participant with deictics, by *ini* first, and in later references by *itu*.

This is similar to the pattern in English in which there is as a rule a two step process of establishing a participant. In English a

participant is introduced using an indefinite article, but thereafter the definite article may be used. For example, consider the beginning of this fairy tale: *'Once upon a time there was a beautiful princess who lived in a large gloomy castle all alone. One day the princess was out in her garden and a strange looking dwarf approached her ...'* In two steps the princess is established in the story. In Bahasa Indonesia we see a three step process of participant introduction and establishment in the discourse involving the use of deictics.

A further interesting characteristic of this textual use of deictics is that it is logically analogous to their spatial use. In the spatial use of deictics objects close to the speaker are referred to as *'this'*, while objects farther away are referred to as *'that'*. Similarly, in the textual use, *ini* *'this'*, occurs on the reference to a participant closest to its introduction (when the hearer first had contact with the participant). *Itu* *'that'*, on the other hand, marks references farther away from the original introduction, maintaining the same close-far away distinction as in the spatial use of deictics.

## 2. SOME VARIATIONS IN THE BASIC PATTERN

While the basic pattern of participant identification described above seems fairly widespread throughout narrative discourse, a survey of a number of narratives soon reveals some seeming contradictions to the pattern. We will examine some of these and show that with an adequate discourse perspective these seeming discrepancies can be rather nicely explained, and further add weight to the growing evidence that a discourse perspective is not a luxury but rather a necessity for linguistic investigation.

Three basic notions of discourse help give a fuller and more complete understanding of the function of deictics in Bahasa Indonesia. These are: 1) a distinction between Major and Minor participants, 2) re-setting of participant identification at a new paragraph, and 3) a distinction between kinds of information in discourse.<sup>4</sup>

We quickly note in observing the use of deictics in Bahasa Indonesia that the pattern described above applies only to major participants (not necessarily people only, for sometimes animals or even inanimate objects play major roles). Minor participants may be introduced in the same fashion with the formula *se-orang ...*, but are subsequently tracked by usually affixing the particle *-nya* rather than one of the deictic words, *ini* or *itu*. For example in the story *'Pengalaman di Jakarta'* a servant who plays a relatively minor role in the whole story

is introduced as *se-orang pelayan hotel* 'one-person servant hotel', but is subsequently tracked by marking him as *pelayannya* 'servant-definite', or when fully established with a pronoun, *dia* 'he'. Thus a different pattern for identifying minor participants exists than that for major participants.

Now consider the following span of identifications of a major participant in the same story, 'Pengalaman di Jakarta':

(29) *pedagang itu*  
*merchant that*

(30) *ia*  
*he*

(31) *dia*  
*he*

(38) *bingungnya,*<sup>5</sup> *ia*  
*shock his he*

(39) *badannya*<sup>5</sup>  
*body his*

(40) *kulitnya*<sup>5</sup>  
*skin his*

(41) *pedagang itu*  
*merchant that*

We may ask why in (41) the reference to the merchant shifts to the fuller noun phrase with *itu* where up to that point pronouns or even more oblique references in the form of third person possessive markers, *-nya*, sufficed. There are no new participants introduced who might be confused with the merchant if pronouns were used. Why then in (41) is there a shift to a full noun phrase, *pedagang itu* 'merchant that', instead of continuing on with a pronoun for reference? The answer lies in the fact that (41) is a paragraph break, the beginning of a new paragraph. This is marked by a time indicator, *akhirnya* 'finally', and a change in location from the shower stall where sentences (29-40) take place to outside the stall. Here as at other paragraph boundaries the reference to a participant is reset more fully, usually by a noun phrase plus the deictic *itu*.

Further refinements in the basic pattern of deictics in participant identification involve a distinction in kinds of information in a discourse. These types of information include setting, background material (additional information given that is not crucial to the advancement of the main event sequence, but which is helpful in explaining the situation so the hearer gets the right information),

evaluative material (comments by the author or speaker giving his evaluation of a situation), and collateral material (information given as to what could have happened, but which did not. This adds much to the events by giving a contrast of what might have occurred) (Grimes 1972:Chapter 4).

In a number of texts observed the basic pattern seemed to fall apart. For example, a sequence of identifications of a major participant, another merchant, goes as follows:

- (55) ... se-orang pedagang yang berasal dari Amerika  
*one-person merchant who originates from America*
- (56) orang Amerika ini  
*person America this*
- (57) kedua pedagang-pedagang itu  
*both merchants that*
- (58) pedagang Amerika ini  
*merchant America this*
- (59) dia  
*he*
- (60) kamar orang Amerika tersebut  
*room person America mentioned*
- (63) orang Amerika ini  
*person America this*
- (64) dia  
*he*
- (65) ia  
*he*
- (66) ia  
*he*
- (67) ia  
*he*
- (68) ia  
*he*
- (70) pedagang Amerika itu  
*merchant America that*

Initially the merchant is introduced according to the basic pattern, se-orang pedagang yang berasal dari Amerika '*one-person merchant who originates from America*'. The second reference is according to the pattern as well, orang Amerika ini '*person America this*'. But then problems arise. Two participants are mentioned, kedua pedagang-pedagang itu '*both merchants that*', and mereka itu '*they that*'. Then again in

(58) the merchant is identified as *pedagang Amerika ini* 'merchant America this', where we might expect *itu* according to the basic pattern described. In (59) things seem semi-normal with a pronominal reference, but then in (60) the merchant's room is identified rather fully as *kamar orang Amerika tersebut* 'room person America mentioned'. Once more in (63) he is *orang Amerika ini* 'person America this', before a shift to a more regular sequence of pronouns occurs, skipping the use of *itu*, leading up to (70) where a fuller noun phrase plus *itu* signals a new paragraph.

This seeming chaos can be ordered and tamed with the help of the concept of kinds of information in discourse. Bahasa Indonesia, like many other languages, often has different grammatical forms used in event passages from those used in non-event material such as setting, background, evaluation of speaker, or collateral.

Such is the case with participant tracking through a discourse. Intervening non-event material within a sequence of events in a sense upsets the normal pattern of participant tracking. In the section described above we see that (57) is background information which the author gives to the readers to help them further understand the connection between the two merchants involved in the story. (58) resumes the event line tracing the initial actions of the American merchant on through (59). In (58) the merchant is identified as *pedagang Amerika ini* and then *dia*, after the interruption of the event line in (57). Once again in (60) through (62) non-event material intervenes, this time setting material, a description of the room and bathroom where the merchant is staying. Within this section, the setting of the stage for the next episode, the room is identified as the merchant's with the rather long phrase *kamar orang Amerika tersebut*.<sup>6</sup> In (63) the sequence of main events resumes with the American merchant being fully identified as *orang Amerika ini*, as a not yet fully established participant, though no other participants have been introduced. Had *itu* been used rather than *ini* it would have marked the merchant as an already well established participant in the discourse. The use of *ini*, however, indicates that the merchant is still not fully established, probably because of the various bits of non-event material being interspersed into the narrative at this point. The interjection of such non-event material in a sense delays the full establishment of the American merchant till well into the episode. By (64) and (65) he is established and referred to with pronouns only.

### 3. CONCLUSIONS

Thus we have seen some important ways in which Bahasa Indonesia uses deictic words to establish a participant in a narrative discourse and help track him through it. Furthermore, it became clear that the kinds of concepts to which it was necessary to appeal to explain choices among deictics in various sentences were not sentence-grammar concepts, but discourse-grammar ones. Thus a discourse perspective is a must in understanding the full function of deictic words in Bahasa Indonesia.

N O T E S

1. This paper was originally presented to the Linguistic Association of the Southwest (LASSO), in Dallas, Texas, November 1976.
2. Or a participant may be named. In such cases the pattern being described does not apply.
3. The scope of this paper is limited to examining primarily the function of the deictic words in participant reference. Tadi and tersebut occur much less frequently and are beyond the scope of this paper.
4. For a full discussion of kinds of information in discourse see Grimes 1972, Chapters 3, 4, and 6.
5. -nya has several functions, two of which are seen in this paper:  
1) as part of the participant identification scheme along with ini and itu, and 2) as the third person possessive morpheme, its function here.
6. I believe the use of tersebut for participant reference is a more direct link to the speech act situation. Its use rather than itu has the effect of drawing the reader/hearer a bit out of the middle of the event sequence, and closer to the speaker-hearer speech act situation. The phrase kamar orang Amerika tersebut '*room person America mentioned*', has a sort of passive feel in the sense of 'the room of the American merchant who was mentioned (by me, the speaker to you, the hearer)'. The performative element is not specified but I feel is implicit in the use of tersebut here.



BIBLIOGRAPHY

GRIMES, Joseph E.

- 1972     *The Thread of Discourse*. Ithaca, N.Y.: Department of  
          Modern Languages and Linguistics, Cornell University.

LONGACRE, Robert E.

- 1976     *An Anatomy of Speech Notions*. Lisse, Belgium: The Peter  
          de Ridder Press.



## QUANTIFIERS IN JAVANESE AND INDONESIAN

MARMO SOEMARMO

### INTRODUCTION

There have been heated discussions for the last ten years or so on the proper analysis of quantifiers, particularly among transformational grammarians. Competing theories use particular analyses to support and/or invalidate basic theoretical claims and assumptions. For instance, generative semanticists such as George Lakoff, Paul Postal, and James McCawley claim that quantifiers are higher predicates. On the other hand, interpretive theorists such as Noam Chomsky and Ray Jackendoff claim that the behaviour of quantifiers is best explained in terms of information available at the surface structures. Recently, Georgette Ioup (1976) attempted to show that grammatical functions (such as subject and direct object) play significant roles in the specification of the scope of quantifiers. Her study can probably be used to support basic theoretical assumptions in relational grammar, a grammatical theory proposed by Paul Postal and David Perlmutter. For an exposition on relational grammar, see the works of Johnson (1974), Lawler (1975), deals with Achenese, and Chung (1976), deals with Indonesian.

This paper deals with various forms and classes of quantifiers in Javanese and Indonesian (Bahasa Indonesia). The focus is on an exposition of the data without any commitment to formal analysis. I will try to present a formal analysis dealing with the syntactic and semantic behaviour of quantifiers in Javanese and Indonesian in the near future. The present work is thus 'theory free', if such situation exists at all. I will, for convenience, occasionally use rewrite rules to summarise and/or generalise certain constructions.

I hope to accomplish two goals: (a) to solicit confirmation, correction, and addition from readers who are working on Javanese, Indonesian, and other Austronesian languages, so that we eventually can accumulate an accurate and coherent picture of the data at hand; and (b) to provide theoretical linguists with theoretically unbiased presentation of data that contain sufficient information so that they will not be misled by a small portion of the data when they are searching for supporting data in these languages. Of course this work cannot claim to be comprehensive or exhaustive, partly due to the nature of quantifiers. They do not belong to one neat class, but across classes. See Akmajian and Lehrer (1976) for an example of a problem in determining the class and relationship of quantifiers and nouns.

Since Indonesian (Bahasa Indonesia) is a national language learned as a second language by most Indonesians (including myself), judgements of well-formedness and degrees of well-formedness vary a great deal, from one regional language speaker to another, and even among the same regional language speakers. Most disagreements centre around what one considers a borrowing may or may not be one for another. Thus, Javanism, Sundaism, and so on are commonly used to explain one's disagreement on the well-formedness of a sentence or construction. I always find it necessary to resort to a regional language such as Javanese in working with Indonesian, although such approach does not guarantee a solid judgement on well-formedness. My training in linguistics makes it impossible for me not to use intuitive knowledge as my source. I think it is safe to say that my intuition about Indonesian is as secure as anybody who speaks the language as a second/national language.

Javanese words and phrases will be written in capital letters, while Indonesian words and phrases will be written in lower case letters. In both cases the new conventional orthography will be used. The new orthography is a uniform orthography adopted by the Indonesian and Malaysian governments in 1973. Ill-formed sentences or forms will be marked with asterisks (\*), questionable ones will be marked with a question mark (?), and a combination of the two (\*?) will be used to mark constructions that are questionable, but more ill-formed than well-formed. Unless otherwise indicated, sentences are based on my own dialect.

## FORMS AND CLASSES

Most works dealing with quantifiers concentrate on the analysis of words and phrases that qualify nouns, such as *'all'*, *'some'* and *'every'*. I will include different types of quantifying words and phrases as much as possible, so that the reader can see the interrelationships between different kinds of quantifying words and phrases to gain an access to some (initial) knowledge of the entire system. It will be up to the reader's theoretical orientation to decide whether the entire system ought to be treated as a unified system or separate sub-systems. Since words in Javanese and Indonesian, and most Austronesian languages, consist of roots (or stems) and affixes, I will present the data on the basis of their forms and make cross references to their classes. The productiveness of derivational processes in these languages provides us with overt relations across different types of quantifying words and phrases, which in other languages may be expressed in completely different forms. In terms of their derivational productiveness, I will start with two major classes to reveal the system, the indefinite quantifiers and the definite quantifiers. The latter contains numerals such as *'two of the men'* and *'two men'*. The rest will be included in the former. The data will be presented in terms of their roots and derivations, one root at a time, followed by a discussion on each form, and a (cumulative) summary of their forms and classes.

## A. INDEFINITE QUANTIFIERS

ROOT 1: AKEH (banyak) - *'many'*

	Javanese	Indonesian	Gloss
1.	AKEH	banyak	<i>'a lot; many; much'</i>
2.	AKEH-AKEH-E	ke-banyak-an	<i>'most; the majority'</i>
3.	KA-AKEH-AN	ke-banyak-an	<i>'too much; overdone'</i>
4.	?SAK-AKEH	se-banyak	<i>'as many/much as'</i>
5.	SAK-AKEH-AKEH-E	se-banyak-banyak-nya	<i>'at the most'</i>
6.	-	ber-banyak-banyak	<i>'in great number'</i>

FORM 1: AKEH (banyak) - *'many'*

The root AKEH (banyak) is normally used as a predicate, as in:

(1) BUKUKU AKEH = AKEH BUKUKU

Buku saya banyak = Banyak buku saya

*'My books are many = (Many are my books)'*

Javanese and Indonesian allow the reversal of subject and predicate, so the basic word order of affirmative simple sentences is either subject-predicate or predicate-subject. There is some semantic consequence (as well as phonological) to this reversal, but we will ignore it for the time being. It is helpful to indicate the two main sentence parts, so I will mark them with a slanted line (/).

When used as a modifier of a noun, Indonesian allows more freedom (in some dialects, but not mine) than Javanese. Consider (2) and (3), and compare the order of Q and N.

- (2) Saya / membeli buku banyak. = Membeli buku banyak / saya.  
       I       bought books many  
       AKU / TUKU BUKU AKEH = TUKU BUKU AKEH / AKU.  
       'I bought many books.'

- (3) Saya / membeli banyak buku. = Membeli buku banyak / saya.  
       AKU / TUKU \*AKEH BUKU. = TUKU \*AKEH BUKU / AKU.  
       'I bought many books.'

Some dialects allow Q+N in Indonesian, in addition to N+Q which is the only order allowed in Javanese. The Q+N order is not in line with the common order of N+Modifier in both languages. For example BUKU LARANG = buku mahal = 'expensive books'.

Because of the availability of Q+N, surface structures of some constructions as in (4) below can be misleading. The Q and the contiguous N may be either an NP or Q is a predicate and N its subject.

- (4) Banyak anak / membeli buku. = \*Membeli buku / banyak anak.  
       children  
       'Many children buy books.'

I can only get a generic interpretation of (4), and reversal seems to result in an ill-formed construction. Compare:

- (5) MANUK / BISA MABUR. = \*BISA MABUR / MANUK.  
       bird can fly  
       Burung / dapat terbang. = \*Dapat terbang / burung.  
       'Birds can fly.'

If a determiner is added to Q+N, then Q+N+Det is an NP which is non-generic. The sentence becomes questionable at best.

- (6) \*?Banyak anak itu / membeli buku. = ?Membeli buku / banyak anak itu.  
       'The many children bought books.'

The situation is complicated by the fact that Javanese allows a special generic sentence with determiner, as in:

- (7) WONG LANANG KUWI / AJA GAMPANG NANGIS.  
 people-male-det IMP-neg - easy - cry  
*'Being a man you should not cry easily.'*

The determiner KUWI seems to have the reading of /non-specific/. See Soemarmo (1970 and 1975) for a discussion on determiners. It is not clear to me at the moment why the reversed structure in (6) is better than the subject-predicate counterpart. It seems that the normal subject position does not take Q+N(DET). If a Q+N(Det) sequence occurs, the Q is a predicate, as in:

- (8) Banyak / anak yang membeli buku. = Anak yang membeli buku / banyak.  
 AKEH / BOCAH SING TUKU BUKU. = BOCAH SING TUKU BUKU / AKEH.  
*'(Many are the children who bought books) = The children who bought books are many.'*
- (9) \*Banyak / anak itu yang membeli buku. =  
 \*Anak itu yang membeli buku / banyak.  
 \*AKEH / BOCAH KUWI SING TUKU BUKU. =  
 \*BOCAH KUWI SING TUKU BUKU / AKEH.

The English translation may be confusing, since English requires *'the'* with relative clause. In Indonesian and Javanese a relative clause makes the head noun specific. Sentence (9) is ill-formed because NP contains Det (itu = KUWI). The only possible interpretation of (9) is when topicalisation is involved. The sentence should be read as follows (double slanted lines (//) after a phrase indicates that the phrase preceding the marker is a topicalised phrase):

- (10) Banyak / anak itu // yang membeli buku. =  
 Anak itu // yang membeli buku / banyak.  
*'The children, the ones who bought books are many.'*

We can trace the structure as follows:

- (11) Anak banyak itu / membeli buku.  
*'The many children bought books.'*  
 topicalise anak itu: Anak itu // banyak / yang membeli buku.  
 reverse: Anak itu // yang membeli buku / banyak.  
 reverse: Banyak / anak itu // yang membeli buku.

As stated above, there is less confusion in Javanese due to the unavailability of Q+N which is an NP.

Indonesian also allows banyak to be used as an adverb of frequency, meaning *'often'*.

- (12) Saya banyak melihat anak yang tidak punya baju didesa itu.  
*I often see children who no have clothes in the village*  
*'I often see children who do not have clothes in the village.'*

Javanese AKEH can not be used as a frequency adverb. There is another word in Indonesian that can be used to express 'often'. The word is sering, and its corresponding Javanese word is SERING or KEREK, both can be used as adverb of frequency.

FORM 2: AKEH-AKEH-E (ke-banyak-an) - 'most'

The reduplicated Q plus a suffix E (or ke+Q+an in Indonesian) can be used to modify a noun.

- (13) ?AKEH-AKEHE BOCAH CILIK / ORA SENENG KOPI.  
 Kebanyakan anak kecil / tidak suka kopi.  
*'Most small children do not like coffee.'*

The Javanese sentence is better expressed with topicalised subject which is also allowed in Indonesian:

- (14) BOCAH CILIK KUWI // AKEH-AKEHE / ORA SENENG KOPI.  
 Anak kecil itu // kebanyakan / tidak suka kopi.  
*'Small children, most of them do not like coffee.'*

Reversal can be applied to (14) to get:

- (15) BOCAH CILIK KUWI // ORA SENENG KOPI / AKEH-AKEHE.  
*'Small children, they don't like coffee, most of them.'*

Second reversal will produce:

- (16) AKEH-AKEHE / BOCAH CILIK KUWI // ORA SENENG KOPI.  
*'Most of them, the small children do not like coffee.'*

Reversal of topic and comment is also allowed to get:

- (17) AKEH-AKEHE / ORA SENENG KOPI // BOCAH CILIK KUWI..  
 Kebanyakan / tidak suka kopi // anak kecil itu.  
*'Most of them, they don't like coffee, those small children.'*

Form 2 cannot be used as a predicate. For Indonesian kebanyakan can be used as a predicate but it has a different meaning, and the form is homophonous with form 3 discussed in the next section.

FORM 3: KA-AKEHAN (ke-banyak-an) - 'too much/many'

The form KA+Q+AN (ke+Q+an) is used as a predicate. In addition to kebanyakan, Indonesian has an alternate form terlalu banyak which some speakers prefer better than kebanyakan.



- (18) SEGANE / KAKEHAN. = KAKEHAN / SEGANE.  
 Nasinya / kebanyakan. = Kebanyakan / nasinya. =  
 Nasinya / terlalu banyak. = Terlalu banyak / nasinya.  
*'The rice is too much. = (Too much is the rice).'*

The form can be used to modify a noun but only to express existentiality or a nominalised adverbialised verb.

- (19) OMAH KUWI / KAKEHAN BOCAH. = KAKEHAN BOCAH / OMAH KUWI.  
*house*  
 Rumah itu / kebanyakan anak. = Kebanyakan anak / rumah itu.  
 ?Rumah itu / terlalu banyak anak. = \*Terlalu banyak anak / rumah  
 itu.  
*'There are too many children in that house.'*

The English translation is misleading or not precise. It seems better to translate the copulative sentences in (19) with copula 'have' in English. So (19) means *'The house has too many children'*. When *terlalu banyak* is used, it seems that *rumah itu* ('the house') has to be a location. Compare the following:

- (20)a. ?Dirumah itu / terlalu banyak anak.  
 b. Dirumah itu / ada terlalu banyak anak.  
 c. Terlalu banyak anak / dirumah itu.  
 d. ?Terlalu banyak anak / ada dirumah itu.

*Di* is a locative marker and *ada* an existential verb. (20a) is questionable because of the deletion of *ada* and *ada* cannot be used in (20d). What happens is that (20) involves two processes: topicalisation and *ada* deletion. Let's trace the structures:

- (21) Basic sentence:  
 (a) Ada/terlalu banyak anak dirumah itu.  
 (predicate) / (object)  
*'There are too many children in the house.'*

Locative topicalisation:

- (b) Dirumah itu // ada/terlalu banyak anak.  
 (see 20b)

*ada* deletion is not desirable in (b), so (20a) is not so good.

Reversal of topic-comment will give:

- (c) ?Ada/terlalu banyak anak / dirumah itu //.

It seems that *ada* deletion is preferred in (c), so (20c) is acceptable.

Note that the only way to get (20d) is by reversing the topic and the object. This process is generally not available, so (20d) is questionable.

The requirement that form 3 must be used in existential NP is clear when we contrast (21) with sentences containing regular non-existential nounphrase, as in (22).

- (22) \*AKU / TUKU KAKEHAN BUKU. = \*TUKU KAKEHAN BUKU / AKU.  
 \*Saya / membeli kebanyakan buku. = \*Membeli kebanyakan buku /  
 saya.  
*'I bought too many books.'*

Indonesian allows a construction like (22) with *terlalu banyak*:

- (23) Saya / membeli terlalu banyak buku. = Membeli terlalu banyak  
 buku / saya.

However, if we resort to Javanese, it seems that *terlalu banyak* modifies the verb instead of the noun. Furthermore, (24) is also available, which clearly shows the adverbial character of *terlalu banyak*.

- (24) Saya / terlalu banyak membeli buku.

In Javanese, the only way to express *'I bought too many books'* is using a nominalised predicate as the subject and the quantifier as a predicate:

- (25) OLEHKU TUKU BUKU / KAKEHAN. = KAKEHAN / OLEHKU TUKU BUKU.  
*'My buying books (is) too many.'*

The predicate describes the quantity of an act, which makes it more of an adverb. If form 3 occurs as a modifier of a noun, it can only modify a nominalised verb, such as:

- (26) BOCAH KUWI / KAKEHAN OMONG.  
 Anak itu / kebanyakan bicara.  
 Anak itu / terlalu banyak bicara.  
*'The children, their talk is too much. (= The children talk too much) = The children did too much talking.'*

FORM 4: ?SAK-AKEH (se-banyak) - *'as many/much as'*

The form SAK (se)+Q is used to express comparative quantity. The Indonesian form is very general. Se is a prefix indicating single, or unit, and se+ADJ generally means *'as+ADJ+as'*, as in *'as big as'*, *'as long as'*, etc. The corresponding Javanese form is not widely acceptable. A more productive form is PADA+ADJ+E KARO to express *'the same ADJ as'*.

- (27) Bukumu itu / sebanyak buku saya. = Sebanyak buku saya / bukumu itu.  
 ?BUKUMU KUWI / SAKAKEH BUKUKU. = ?SAKAKEH BUKUKU / BUKUMU KUWI.  
*'Your books are as many as my books.'*

Javanese prefers (28) to express (27):

- (28) BUKUMU KUWI // PADA AKEHE KARO BUKUKU. = PADA AKEHE KARO BUKUKU / BUKUMU KUWI.  
*'Your books are the same amount as my books.'*

Sentence (28) comes from:

- (29) AKEHE BUKUMU KUWI / PADA KARO (AKEHE) BUKUKU.  
*'The amount of your books is the same as (the amount) of my books.'*

To get (28), topicalise BUKUMU KUWI in (29). A second topicalisation is possible. Topicalise AKEHE in (28) and you get:

- (30) BUKUMU KUWI // AKEHE // PADA KARO BUKUKU.  
*'Your books, their amount, it's the same as my books.'*

FORM 5: SAK-AKEH-AKEH-E (se-banyak-banyak-nya) - 'at the most'

The form SAK (se) + reduplicated Q + E (nya) is a productive form. In general it means 'as-ADJ-as possible', as in 'as big as possible', 'as many as possible', etc. It is used mostly as a predicate, as in:

- (31) TUKUA SAKAKEH-AKEHE. = Belilah sebanyak-banyaknya.  
*'Buy as many as possible.'*

When used to modify a numeral, it acts as a limiting adjective, the equivalent of 'at the most' in English. For instance:

- (32) WONG DESA KUWI / SAKAKEH-AKEHE SEWU.  
 Orang desa itu / sebanyak-banyaknya seribu.  
*'The villagers are at the most one thousand.'*

Semantically, the surface structure (32) is misleading. It seems that the sentence contains a deleted element 'the amount' or 'the number'.

(32) is synonymous with (33):

- (33) Jumlah orang desa itu / sebanyak-banyaknya seribu.  
*'The number/amount of the villagers is at the most 1000.'*

Furthermore, constructions such as (34) are also available:

- (34) SAKAKEH-AKEHE WONG DESA KUWI / SEWU.  
 Sebanyak-banyaknya orang desa itu / seribu.  
*'The most number of the villagers is one thousand.'*

(32) seems to have a structure that contains topicalisation of *'the villagers'*, which can be represented as follows:

- (35) WONG DESA KUWI // SAKAKEH-AKEHE / SEWU.  
 Orang desa itu // sebanyak-banyaknya / seribu.  
*'Those villagers, their largest amount is 1000.'*

Somewhat similar to the use of form 5 in (34) is in sentences such as:

- (36) SAKAKEH-AKEHE BUKUMU / DURUNG SEPIRAA YEN DIBANDHING KARO BUKUNE  
 ALI. = Sebanyak-banyaknya bukumu itu / belum seberapa kalau  
 dibandingkan dengan buku Ali.  
*'No matter how much/many your books are, it's nothing compared  
 to Ali's.'*

The above construction is very productive. In general, the form can be used to express *'no matter how ADJ ...'* as in *'no matter how big ...'*, *'no matter how expensive ...'*, etc.

FORM 6: ber-banyak-banyak - *'in large groups'*

The form *ber* + reduplicated Q does not have a Javanese counterpart and does not occur in my dialect. However, some speakers use expressions such as:

- (37) Mereka datang berbanyak-banyak.  
*'They came in large groups.'*

The form is intuitively possible, since forms such as *ber-bondong-bondong* *'in groups'* are available.

FORM 7: AKEH-AKEH ( banyak-banyak) - *'so much'*

The complete reduplication of the root is not a productive process. Its limited use can be illustrated by sentences such as:

- (37a) AKEH-AKEH AREP KOK ANGGO APA?  
 ?Banyak-banyak akan kamu pakai apa?  
*'What are you going to do with so much?'*

## ROOT 2: SETHITHIK (sedikit) - 'a little, a few'

Javanese	Indonesian	Gloss
1. SETHITHIK	sedikit	'a few, a little'
2. SETHITHIK-2-E	?sedikit-2-nya	'at least'
3. KE-SETHITHIK-EN	*ke-sedikit-an	'too few'
4. *SE-SETHITHIK	?se-sedikit	'as few as'
5. SAK-SETHITHIK-2-E	?se-sedikit-2-nya	'as few as possible'
6. -	?ber-sedikit-2	'in small groups' (?)
7. SETHITHIK-2	sedikit-sedikit	'with the least provocation'
8. SETHITHIK-2	sedikit (demi)sedikit	'bit by bit'
9. ORA SETHITHIK-2-A	tidak sedikitpun	'not the least bit'

## FORM 1: SETHITHIK (sedikit) - 'a few, a little, a small quantity'

The root form can be used as a predicate, as in:

- (38) BUKUKU / SETHITHIK. = SETHITHIK / BUKUKU.  
 Buku saya / sedikit. = Sedikit / buku saya.  
 'My books are a few. = (A few are my books).'

As is the case with AKEH (banyak), when used as a noun modifier, some Indonesian dialects allow the alternate Q+N order:

- (39) AKU / TUKU BUKU SETHITHIK. = TUKU BUKU SETHITHIK / AKU.  
 Say / membeli buku sedikit. = Membeli buku sedikit / saya.  
 'I bought a few books.'
- (40) \*AKU / TUKU SETHITHIK BUKU. = \*TUKU SETHITHIK BUKU / AKU.  
 ?Saya / membeli sedikit buku. = ?Membeli sedikit buku / saya.  
 'I bought a few books.'

The Q+N order with sedikit seems worse (at least in my dialect) than the one with banyak. It is even worse when Q+N is in subject position:

- (41) \*?Sedikit anak / membeli buku. = \*Membeli buku / sedikit anak.  
 'A few children buy books.'

I can only get a generic interpretation, if at all, and the reversed order is completely ill-formed. The Q+N+Det is also ill-formed:

- (42) \*Sedikit anak itu / membeli buku. = \*Membeli buku / sedikit anak itu.  
 (intended to be: 'The few children bought books')

The N+Q(Det) forms behave exactly the same as AKEH (banyak):

- (43) BOCAH SETHITHIK / ORA DADI PIKIRAN. = ?ORA DADI PIKIRAN / BOCAH SETHITHIK.

Anak sedikit / tidak jadi beban. = ?Tidak jadi beban / anak sedikit.

*'A few children will not be a burden.'* (generic)

- (44) BOCAH SETHITHIK KUWI // ORA DADI PIKIRAN.

Anak sedikit itu // tidak jadi beban.

*'A few children, (they) are not a burden.'*

- (45) SETHITHIK / BOCAH SING TUKU BUKU KUWI. = BOCAH SING TUKU BUKU KUWI / SETHITHIK.

Sedikit / anak yang membeli buku itu. = Anak yang membeli buku itu / sedikit.

*'(A few are the children who bought the books). = The children who bought books are a few.'*

- (46) \*SETHITHIK / BOCAH KUWI SING TUKU BUKU.

\*Sedikit / anak itu yang membeli buku.

The following sentences involve topicalisation. (47) is the basic sentence (untopicalised):

- (47) BOCAH SETHITHIK KUWI / TUKU BUKU.

Anak sedikit itu / membeli buku.

*'The few children bought books.'*

- (48) BOCAH KUWI // SETHITHIK / SING TUKU BUKU.

Anak itu // sedikit / yang membeli buku.

*'The children, a few are the ones who bought books.'*

- (49) BOCAH KUWI // SING TUKU BUKU / SETHITHIK.

Anak itu // yang membeli buku / sedikit.

*'The children, the ones who bought books are a few.'*

- (50) SETHITHIK / BOCAH KUWI // SING TUKU BUKU.

Sedikit / anak itu // yang membeli buku.

*'A few, the children who bought books.'* (or better, literal translation: *'A few, the children, those who bought books.'*)

While banyak can be used as frequency adverb meaning 'often', sedikit does not seem to be available to express 'rarely':

- (51) \*saya sedikit membaca buku roman.

*'I rarely read a novel.'*

FORM 2: SETHITHIK-SETHITHIK-E (?sedikit-sedikit-nya) - 'at least'

The reduplicated Q plus a suffix E (nya) with SETHITHIK (sedikit) breaks the pattern with the same form with AKEH (banyak) as Q. Compare (52) with (13):

- (52) \*SETHITHIK-SETHITHIK-E BOCAH CILIK / ORA SENENG KOPI.  
 \*Sedikit-sedikit-nya anak kecil / tidak suka kopi.  
 (cannot be used to express the opposite of 'Most small children do not like coffee')

Similarly, if SETHITHIK-SETHITHIK-E (sedikit-sedikit-nya) is used to replace sentences with AKEH-AKEH-E (ke-banyak-an) in (14)-(17), the sentences with SETHITHIK (sedikit) roots will be ill-formed. Ke-sedikit-nya does not occur in Indonesian.

What is interesting about SETHITHIK-SETHITHIK-E (sedikit-sedikit-nya) is that semantically, the forms are the opposite of SAK-AKEH-AKEH-E (se-banyak-banyak-nya). In other words, the SE- (se-) is considered a prefix corresponding to SAK (se). Compare the following sentences with sentences (31)-(36).

- (53) TUKUA / SETHITHIK-(SE)THITHIK-E.  
 Belilah / sedikit-(se)dikit-nya.  
 'Buy as little as possible.'

The forms in (53) seem to come from:

- (54) TUKUA / SAK-SETHITHIK-SETHITHIK-E.  
 Belilah / se-sedikit-sedikit-nya.

and there seems to be a phonological (surface) rule that prevents identical (closely similar) syllables in succession. The corresponding sentence (32) below is an interesting transition. The Javanese sentence allows both forms, but not the Indonesian one:

- (55) WONG DESA KUWI / SAK-SETHITHIK-SETHITHIK-E SEWU. =  
 WONG DESA KUWI / SETHITHIK-SETHITHIK-E SEWU.  
 Orang desa itu / \*se-sedikit-sedikit-nya seribu. =  
 Orang desa itu / sedikit-sedikit-nya seribu.  
 'The villagers are at least one thousand.'

Some Indonesian dialects allow further reduction of sedikit-sedikit-nya into sedikit-dikit-nya.

As is the case with AKEH (banyak), in sentences such as (56) seem to imply the deletion of 'the amount':

- (56) Orang desa itu / sedikit-sedikitnya seribu. =  
Jumlah orang desa itu / sedikit-sedikitnya seribu.  
 (amount/total number)  
*'The number of villagers is at least one thousand.'*

Although Javanese has the corresponding sentence to the first sentence in (56), it does not have the corresponding word for jumlah, the closest is AKEHE which corresponds to Indonesian banyaknya. Some speakers, however, do not accept banyaknya.

- (57) ?Banyaknya orang desa itu / sedikit-sedikitnya seribu.  
 AKEHE WONG DESA KUWI / (SAK)-SETHITHIK-SETHITHIK-E SEWU.  
*'The number/amount of villagers is at least one thousand.'*

The form can also be used to express 'no matter how ...':

- (58) SAK-SETHITHIK-SETHITHIKE WONG DESA KUWI / KUDU KOK GATEKAKE.  
 ?Sedikit-sedikitnya orang desa itu / harus kau perhatikan.  
*'No matter how few the villagers are you must pay attention to them.'*

The Indonesian sentence does not seem well-formed. A more acceptable expression to express the same meaning is:

- (59) Meskipun (jumlah) penduduk didesa itu (hanya) sedikit, kamu harus memperhatikan mereka.  
*'Even though the (number of) people in the village is (only) a few, you must pay attention to them.'*

FORM 3: KE-SETHITHIK-EN (\*ke-sedikit-an) - 'too few'

There is a gap in Indonesian in the form ke+Q+an. While we have ke+banyak+an meaning 'too much' there is no \*ke+sedikit+an to express 'too few'. Javanese, on the other hand, has KE-SETHITHIK-EN meaning 'too few' as the opposite of KE-AKEH-AN (= KAKEHAN) meaning 'too many'. It is not surprising that some Javanese would use ke-sedikit-an when they speak Indonesian.

The primary occurrence of this form is as a predicate:

- (60) SEGANE / KESETHITHIKEN. = KESETHITHIKEN / SEGANE.  
 Nasinya / terlalu sedikit. = Terlalu sedikit / nasinya.  
*'The rice is too little.'*

When used to modify a noun, the modified noun expresses existentiality:

- (61) KOPI KUWI / KESETHITHIKEN GULA.  
 ?Kopi itu / terlalu sedikit gula.  
*'There is too little sugar in the coffee.'* (or, literally:  
*'The coffee, there is too little sugar (in it).'*)



Most speakers prefer the predicate of the form, and express (61) as:

- (62) KOPI KUWI // GULANE / KESETHITHIKEN.  
 Kopi itu // gulanya / terlalu sedikit.  
*'The coffee, its sugar is too little.'*

Indonesian allows the form to be used in object position modifying the direct object to express sentences such as *'I bought too little sugar'*. However, Javanese does not allow such form.

- (63) Saya / membeli terlalu sedikit gula.  
 \*AKU / TUKU KESETHITHIKEN GULA.  
*'I bought too little sugar.'*

Javanese expresses (63) with a nominalised predicate:

- (64) OLEHKU TUKU BUKU / KESETHITHIKEN.  
*'My buying the books is too little.'*

The English gloss for (64) may be misleading or not clear. The nominalised form OLEHKU TUKU BUKU expresses the presupposition AKU TUKU BUKU *'I bought books'*. The predicate KESETHITHIKEN describes the amount of books in the presupposition. The predicate may also describe the predicate of the presupposition, as in:

- (65) OLEHKU TUKU BUKU / KEKEREPEN.  
*'My buying books is too often. = I bought books too often.'*

For a possible explanation to the nominalised requirement, see Soemarmo 1977.

The form can be used to modify a nominalised verb, if Q is AKEH, as in (26) above, but it does not seem well-formed with SETHITHIK:

- (66) ?BOCAH KUWI / KESETHITHIKEN OMONG.  
 Anak itu / terlalu sedikit bicara.  
*'The child (has) too little talk. (= The child talks too little.)'*

A more acceptable way of expressing (66) in Javanese is by topicalisation:

- (67) BOCAH KUWI // OMONGE / KESETHITHIKEN.  
*'The child, its talk is too much.'*

Sentence (67) comes from:

- (68) OMONGE BOCAH KUWI / KESETHITHIKEN.  
*'The child's talk is too little.'*

Possessive constructions are formed by NP+E+NP, so OMONG is the possessed noun, E the connector, and BOCAH KUWI the possessor.

FORM 4: \*SE-SETHITHIK (?se-sedikit) - *'as few/little as'*

In Indonesian se-Adjective is a productive form to express comparative *'as ... as'*, but Javanese does not use SE+Adjective. Instead, it makes use of the form PADA+Adjective+E. However, neither form seems to accept SETHITHIK (sedikit).

## (69) \*BUKUKU SE-SETHITHIK BUKUMU.

?Buku saya se-sedikit bukumu.

*'My books are as few as your books.'*

## (70) ?BUKUKU PADA SETHITHIK-E KARO BUKUMU.

A better way to express (70) is by a conjunction:

## (71) BUKUKU KARO BUKUMU KUWI PADA SETHITHIKE.

*'My books and your books are of the same (fewness?).'*

The structure in (71) is productively used to express comparison *'the same Adjective-ness'*. The conjunction is formed in a manner similar to the formation of English sentences such as:

(72) *'My books are similar to your books.'* + *'My books and your books are similar.'*FORM 5: SAK-SETHITHIK-SETHITHIK-E (?se-sedikit-sedikit-nya) - *'as few as possible'*

In discussing form 2, we noted that form 5 can be reduced to form 2 by phonological surface redundancy rule. The uses of form 5 are discussed above.

FORM 6: ?ber-sedikit-sedikit - *'in small groups'*

The form ber+Q occurs only in Indonesian. Ber-banyak-banyak is used by some speakers to express *'in large groups'*, but it does not seem generalisable. Ber-sedikit-sedikit does not seem to be well-formed.

FORM 7: SETHITHIK-SETHITHIK (sedikit-sedikit) - *'with the least provocation'*

The simple reduplication of Q is available with SETHITHIK (sedikit), but not with AKEH (banyak). It is used as an adverb of some sort, similar to frequency adverbs such as *'often'* or *'rarely'*.

## (73) BOCAH KUWI / SETHITHIK-SETHITHIK NANGIS.

?Anak itu / sedikit-sedikit menangis.

*'The child cries with the least provocation.'*

FORM 8: SETHITHIK-SETHITHIK (sedikit (demi) sedikit) - '*bit by bit*'

The simple reduplicated Q can also be used as manner adverb meaning '*bit by bit*'. In Javanese, another form, KANTHI SETHITHIK is also available, and in Indonesian *demi* is normally used to connect the two identical Qs.

(74) \*?BOCAH KUWI / MANGAN SEGA SETHITHIK-SETHITHIK.

Anak itu / makan nasi sedikit demi sedikit.

*'The child ate the rice bit by bit.'*

As stated above, some adverbs are expressed by topicalised construction in Javanese, and for manner adverb, nominalisation of the predicate is necessary. A better way to express (74) in Javanese is:

(75) BOCAH KUWI // OLEHE MANGAN SEGA / SETHITHIK-SETHITHIK. =

BOCAH KUWI // MANGANE SEGA / SETHITHIK-SETHITHIK.

*'The child, his manner of eating rice is bit by bit. =*

*The child, his eating rice is bit by bit.'*

The first construction is used more often than the second. Sentences in (75) come from (76) and (77), respectively.

(76) OLEHE MANGAN SEGA BOCAH KUWI / SETHITHIK-SETHITHIK.

*'The child's manner of eating rice is bit by bit.'*

(77) MANGANE SEGA BOCAH KUWI / SETHITHIK-SETHITHIK.

*'The child's eating rice is bit by bit.'*

Form 7 can be distinguished from form 8 by the availability of form 8 in topicalised construction. Form 7 cannot occur in topicalised constructions similar to (76) or (77).

(78) \*BOCAH KUWI // (OLEHE) NANGISE / SETHITHIK-SETHITHIK.

cannot mean *'The child cries with the least provocation.'*

FORM 9: ORA SETHITHIK-SETHITHIK-A (tidak sedikit-pun) - '*not the least bit*'

The form reduplicated Q + A (pun) can only be used in negative form. The form cannot be used with root 1, AKEH (banyak), but it can be used with other limited number of Qs and adjectives.

(79) BOCAH KUWI ORA MANGAN SETHITHIK-SETHITHIKA.

Anak itu tidak makan sedikitpun.

*'The child did not eat at all.'*

In constructions where topicalisation is allowed, the negative ORA (tidak) must go together with reduplication Q + A (pun).

- (80) Tidak sedikitpun // anak itu / diberi nasi.  
 ?ORA SETHITHIK-SETHITHIK-A // BOCAH KUWI / DIWENEHI SEGA.  
 (?) *'Not even a little bit was the child given rice.'*

In any case, whether the negative is immediately following Q or not, the negative morpheme must be there.

- (81) \*BOCAH KUWI / MANGAN SETHITHIK-SETHITHIK-A.  
 \*Anak itu / makan sedikitpun.

Topicalised sentence (80) comes from:

- (82) Anak itu / tidak diberi nasi sedikitpun.  
*'The child was not given rice at all (not even a little).'*

The subject of (82) can be topicalised to get:

- (83) Anak itu // tidak sedikitpun (dia) diberi nasi.  
*'The child, not even a little was he given rice.'* (?)

#### SUMMARY

Before we include other roots, it may be useful at this juncture to compare the formation and classes derived from roots 1 and 2.

(a) Semantically, root 1 is the opposite of root 2, so the negated form of one is synonymous with the positive form of the other. For example, ORA AKEH (tidak banyak) is the opposite of SETHITHIK (sedikit).

(b) In terms of their morphological processes, there are shared processes across roots in the two languages. The charts below display the similarities and differences of the two roots.

Formation	Root 1	Root 2
1. UNAFFIXED Unaffixed	AKEH banyak	SETHITHIK sedikit
2. ROOT-ROOT-E (root-root-nya) ke-root-an	AKEH-AKEH-E *banyak-banyak-nya ke-banyak-an	SETHITHIK-SETHITHIK-E sedikit-sedikit-nya *ke-sedikit-an
3. KE-root-EN ke-root-an	KE-AKEH-EN ke-banyak-an	KE-SETHITHIK-EN *ke-sedikit-an
4. (SE-ROOT) PADA-ROOT-E se-root	*SE-AKEH PADA-AKEH-E se-banyak	*SE-SETHITHIK PADA-SETHITHIK-E se-sedikit
5. SAK-ROOT-ROOT-E se-root-root-nya	SAK-AKEH-AKEH-E se-banyak-banyak-nya	SAK-SETHITHIK-SETHITHIK-E *se-sedikit-sedikit-nya
6. - ber-root-root	- ber-banyak-banyak	- *ber-sedikit-sedikit

Formation	Root 1	Root 2
7. ROOT-ROOT root-root	AKEH-AKEH ?banyak-banyak	SETHITHIK-SETHITHIK sedikit-sedikit
8. ROOT-ROOT root (demi) root	*AKEH-AKEH *banyak demi banyak	SETHITHIK-SETHITHIK sedikit demi sedikit
9. NEG-ROOT-ROOT-A NEG-root-pun	*ORA AKEH-AKEH-A *tidak banyak-pun	ORA SETHITHIK-SETHITHIK-A tidak sedikit-pun

The forms above constitute different classes:

Formation	Possible classes with Root 1	Possible classes with Root 2
1. UNAFFIXED Unaffixed	PRED; NQ pred; nq; adv	PRED; NQ pred; nq
2. ROOT-ROOT-E ke-root-nya	NQ nq	NQ -
3. KE-ROOT-EN ke-root-en	PRED; N'Q; NOM.Q pred; n'q	PRED; N'Q; NOM.Q -
4. PADA-ROOT-E se-root	COMPAR. compar.	COMPAR. -
5. SAK-ROOT-ROOT-E se-root-root-nya	NP/INFINITE AMOUNT np/infinite amount	NQ nq
6. - ber-root-root	- nq	- -
7. ROOT-ROOT root-root	INFINITE AMOUNT ?infinite amount	ADV-FREQUENCY adv-frequency
8. ROOT-ROOT root-demi-root	- -	ADV-MANNER adv-manner
9. NEG-ROOT-ROOT-A NEG-root-pun	- -	INDEFINITELY SMALL/NP indefinitely small/np

ROOT 3: PIRA (berapa) - 'some indefinite amount'

To facilitate comparison with the two roots presented above, the nine processes above will be used as a frame of reference.

Javanese	Indonesian	Gloss
1. PIRA	berapa	'indefinite/unknown amount'
2. PIRA-PIRA-NE	*berapa-berapa-nya	'the amount'
3. *KE-PIRA-(N)EN	*ke-berapa-an	-
4. ORA SE-PIRA (A)	tidak se-berapa	'not much'
5. ?SAK-SE-PIRA-PIRA-NE	*se-berapa-berapa-nya	'no matter what amount'



- (87) ALI KUWI // OLEHE TUKU BUKU / PIRA? = OLEHE TUKU BUKU / ALI KUWI // PIRA? = PIRA / ALI KUWI // OLEHE TUKU BUKU?  
*'That Ali, how many/much is (his buying books)?'*
- (88) PIRA / BUKU SING DITUKU ALI KUWI? = BUKU SING DITUKU ALI KUWI / PIRA?  
 Berapa / buku yang dibeli Ali itu? = Buku yang dibeli Ali itu / berapa?  
*'How many are the books which Ali bought?'*
- (89) \*ALI / TUKU BUKU PIRA REGANE?  
 \*Ali / membeli buku berapa harganya?  
 (lit. *'Ali bought books for what price/how much?'*)
- (90) PIRA / REGANE BUKU SING DITUKU ALI KUWI?  
 Berapa / harga(nya) buku yang dibeli Ali (itu)?  
*'What is the price of the book(s) which was bought by Ali?'*

In general, questions with question words are formed by making a statement with the questioned part expressed by a pro-form. For example, compare:

- (91) AKU MANGAN SEGA.            *'I ate rice.'*  
 AKU MANGAN APA?            *'(I ate what?) = What did I eat?'*  
 AKU MANGAN MAU ESUK.       *'I ate this morning.'*  
 AKU MANGAN KAPAN?       *'(I ate when?) = When did I eat?'*

Following the general rule for the formation of interrogatives with question words, (86) is available. In Javanese, however, constructions like (86) with PIRA are used more as echo questions, although they are also used as true questions. Javanese makes use of presuppositional expressions, such that questions normally require indication of presupposition (see Soemarmo 1977 for more on this). Sentences such as (87) are used more than (88). You must nominalise the predicate. OLEHE TUKU BUKU *'his buying books'* is the nominal form of TUKU BUKU *'bought books'*. Another common construction is where PIRA is used as an attribute of an NP, as in (85) and (88). With REGANE (*harganya*) *'its price'*, the construction PIRA (*berapa*) + NP seems required. The general formation of interrogatives with question words does not apply to PIRA REGANE (*berapa harganya*), since (89) is ill-formed. It has to be converted to (90). Only this last requirement applies to both languages. The indication of presupposed expression is not required in Indonesian.

FORM 2: PIRA-PIRA-NE (?\*berapa-berapa-nya) - *'the exact amount'*

The reduplication of the root plus E (nya) has a limited use. The N is added by a general rule: if E is added to a word that ends in a vowel, N is inserted. Indonesian does not seem to have this form.

(92) AKU / ORA NGERTI PIRA-PIRA-NE.

*'I don't know the exact/actual amount.'*

(93) AJA MBGK KANDAKAKE PIRA-PIRANE.

*'Don't tell (anybody) how much it is/how many they are.'*

I am not sure whether the negative morpheme is a requirement. The above sentences do not seem to be well-formed without negative morphemes.

(94) ?AKU / NGERTI PIRA-PIRA-NE.

*'I know how much it is/how many they are.'*

?KANDANANA / PIRA-PIRA-NE.

*'Tell (someone) how much it is/how many they are.'*

FORM 3: \*KE-PIRA-NEN (\*ke-berapa-an)

The form KE-Q-EN (ke-Q-an) generally designates an excessive attribute/quality/quantity, such as *'too big'*, *'too much'*, and *'too high'*. It seems semantically impossible to talk about *'an excessive indefinite amount, size, etc.'*. The availability of the form seems to be prevented by a general semantic constraint.

FORM 4: ORA SE-PIRA-(A) (tidak se-berapa) - *'not that much/many'*

The form SE (se) -Root 3- (A) must be used with a negative morpheme. It is used as a predicate, and the subject must be a particular noun-phrase or nominal phrase.

(95) \*OMAHE / SE-PIRA(A).

\*Rumahnya / se-berapa.

(ill-formed without negative)

(96) OMAHE / ORA SE-PIRA(A).

Rumahnya / tidak se-berapa.

*'His/her house isn't that much.'*

The meaning of *'that much'* can be *'having low value'* or *'having low (aesthetic) quality'*. In Javanese, the suffix A can be dropped, probably due to its phonological shape, which is identical to the last sound in the root.



The requirement of NP or nominal subject can be illustrated by sentences such as:

- (97) OLEHE TUKU BUKU / ORA SEPIRA(A).  
*'He didn't buy that many books.'* (lit. *his buying books is not that many.*)
- (98) Buku yang dibelinya / tidak se-berapa.  
 BUKU SING DITUKU / ORA SEPIRA(A).  
*'The books which he bought aren't that many/didn't cost that much.'*

Only Javanese has and makes use of nominalised predicate such as in (97). Both languages use relative clause constructions as in (98). Even though the form is used to specify the quantity or quality of the object, it can not directly modify the NP-object.

- (99) \*ALI / TUKU ORA SE-PIRA(A) BUKU.  
 \*Ali / membeli tidak se-berapa buku.  
*'Ali bought not that many books.'*

FORM 5: ?SAK-SE-PIRA-PIRA-NE (\*se-berapa-berapa-nya)

The form SAK (se) -reduplicated root- NE (nya) does not seem to take PIRA (berapa). Indonesian definitely does not have this form. The Javanese form, if it is available at all, seems to be a variant of form 2 (PIRA-PIRA-NE).

- (100) AKU ORA NGERTI SAK-SE-PIRA-PIRA-NE. = AKU ORA NGERTI PIRA-PIRA-NE.  
*'I don't know the actual amount.'*

While PIRA-PIRA-NE is ambiguous, it can mean *'the actual amount'* or *'the actual cost/price'*, SAK-SE-PIRA-PIRA-NE seems to have only the first meaning.

This form may also be used by some to express what form 7 (discussed below) is expressing.

FORM 6: (\*ber-berapa-berapa)

The form ber+reduplicated root is used with limited number of Q, such as banyak to express *'in ... group'* such as *'in a large group'*. As is the case with form 3, this form is prevented by a general semantic constraint.

FORM 7: PIRA-PIRA(A) (\*berapa-berapa) - *'whatever amount/cost'*

The Javanese forms 2, 5, 7, and 9 have some usage in common, and so in some cases they are interchangeable. It is definitely true that the

reduplicated form is not available in Indonesian, but form 9 (berapa+pun) may fit in either forms mentioned above. I have not been able to sort out the subtle distinctions of these forms. It may take more research involving complex constructions to reveal the difference. The uses of PIRA-PIRA (sometimes the suffix A can be added) can be illustrated by the following:

- (101) PIRA-PIRA (A) TUKUNEN.  
*'No matter how many/much, buy it.'*
- (102) BOCAH KUWI / ORA PIRA-PIRA.  
*'The child does not care/has no mercy.' (i.e. ignores the amount of suffering as a consequence of his act)*
- (103) MBOK SING PIRA-PIRA.  
*'Show some consideration. / Be considerate.' (i.e. beware of the amount of suffering your action may result in)*
- (104) \*TUKUA PIRA-PIRA BUKU.  
*'Buy any amount of books.'*

As is the case with form 2, this form cannot be used to modify a noun, or used as a noun modifier.

FORM 8: ORA PIRA-PIRA (\*tidak berapa-berapa) - *'without mercy'*

The form negative+reduplicated root is used as an idiom whose meaning can be derived from PIRA. See sentence (102) above for an illustration. This form does not occur in Indonesian.

FORM 9: ?SE-PIRA-PIRA-A (berapa-pun) - *'regardless of the amount'*

The form SE+reduplicated root+A (root+pun) is used to express unlimited large/small amount of cost. It can be used as a predicate only, not a modifier.

- (105) SE-PIRA-PIRA-A TAMPANANA.  
*'No matter how big/many/much it costs, take it.'*

To get a better feeling about the meaning of the form, consider the productive derivation using SE + NP:

- (106) SE-CANGKIR     *'a cup'*  
           SE-BUNGKUS   *'a bag (wrapper)'*  
           SE-KILO       *'one kilogram'*

So, SE-PIRA-PIRA is used to refer to indefinite quantity in terms of measurements such as *'cup'*, *'bag'*, *'kilogram'*, etc. In other words,

PIRA-PIRA is the pro-form of measurement NP. The Indonesian *berapa-pun* is used in a similar way:

- (107) *Berapa-pun terimalah.*  
*'No matter how many/much, take it.'*

The reduplicated forms of many roots are used to express pro-forms. For example:

- (108)a. KOWE / TUKU APA?  
 Kamu / membeli apa?  
 (lit. 'You bought what?')
- b. ORA APA-APA.  
 Tidak apa-apa.  
 (lit. 'not anything')
- (109)a. KOWE / AREP NYANG ENDI?  
 Kamu / akan pergi kemana?  
 (lit. 'You will go where?')
- b. ORA NYANG ENDI-ENDI.  
 Tidak kemana-kemana.  
 (lit. 'not anywhere')
- (110)a. KOWE / AREP METHUK SAPA?  
 Kamu / akan menjemput siapa?  
*'You will pick up who?'*
- b. DUDU SAPA-SAPA.  
 Bukan siapa-siapa.  
*'Nobody.'* (lit. *not anybody*)

FORM 10: (be-berapa) - 'some'

Javanese does not have a word to express 'some' as in 'some books'. Indonesian has the form 'partial reduplication of root' *be-berapa*. The corresponding Javanese \*PI-PIRA does not occur. This form is used mostly as noun modifier.

- (111) Saya / membeli beberapa buku.  
*'I bought some books.'*

Unlike root 1 (*banyak* 'many'), this form can only occur preceding the modified noun. The reverse order is ill-formed:

- (112) \*Saya / membeli buku beberapa.

The form can be combined with some words that can combine with numerals (see the section on Definite Quantifiers below), such as:

(113) Saya / makan tiga kali sehari.

'I eat three times a day.'

Saya / makan beberapa kali sehari.

'I eat some/many times a day.'

(114) Saya / membeli tiga buah pisang.

'I bought three bananas.

Saya / membeli beberapa buah pisang.

'I bought some bananas.'

Words such as buah in (114) are called Classifiers. They will be discussed separately below.

Beberapa can be used without classifiers, but the deletion of qualifiers seems freer with countable than uncountable nouns.

(115)a. Saya membeli beberapa buah pisang.

b. Saya membeli beberapa kaleng minyak. ('cans of oil')

c. Saya membeli beberapa pisang.

d. ?\*Saya membeli beberapa minyak.

FORM 11: PIRA-NG-NP (berapa-NP) - 'how many/much NP'

The root can be combined with a quantifier, adjective, or words denoting a measurement. A linking sound NG is added in Javanese to form the compound.

(116) KOWE / AREP LUNGA PIRANG DINA?

Kamu / akan pergi berapa hari?

'You will go for how many days?'

(117) PIRA AKEHE BOCAH KUWI?

Berapa banyak(nya) anak itu?

'How many are the children?'

(118) (SE)PIRA GEDHENE GUNUNG KUWI?

(?Se)berapa besar gunung itu?

'How big is the mountain?'

FORM 12: PIRA-NG-PIRA-NG (banyak) - 'many'

The Javanese form ROOT-NG-ROOT-NG is synonymous with root 1 (AKEH), but this form seems to imply 'an unexpected large amount' or 'an unusually large amount'. Indonesian does not seem to make the distinction.

(119) ANAKE / PIRANG-PIRANG.

ANAKE / AKEH.

*'His/her children are many.'*

The first sentence in (119), unlike the second, is not a matter-of-fact statement. It is either derogatory or an indication of surprise. Derogatory and surprise do not seem to go well with command or request. In other words, (120) is ill-formed because one cannot talk about *'request surprisingly'*.

(120) \*TUKUA SING PIRANG-PIRANG.

Compare (120) with:

(121) TUKUA SING AKEH.

*'Buy many.'*

I am not sure whether English expressions *'many'* and *'plenty'* have a similar distinction as AKEH and PIRANG-PIRANG.

FORM 13: PIRA-NG-PIRA-NG-NP (ber-NP-NP) - *'plenty of NP'*

Form 12 can be combined with NP that can be combined with form 11, in Javanese. Indonesian makes use of ber-reduplicated NP.

(122)a. BUKUKU / PIRANG-PIRANG.

Buku saya / banyak.

*'My books are plenty.'*

b. BUKUKU / PIRANG-PIRANG KRANJANG.

Buku saya / ber-keranjang-keranjang.

*'My books are many baskets. (= I have baskets of books).'*

Just as AKEH is not completely synonymous with PIRANG-PIRANG, the Indonesian banyak does not imply *'excessive amount'*. Note the word order in (122a) in Javanese. When used as a predicate expressing quantity, the order is Q+NP. However, when Q is used as a noun modifier, the order is the reverse:

(123)a. AKU / TUKU BUKU PIRANG-PIRANG.

*'I bought plenty of books.'*

b. \*AKU / TUKU PIRANG-PIRANG BUKU.

c. AKU / TUKU BUKU PIRANG-PIRANG KRANJANG.

*'I bought baskets of books.'*

The form does not seem to behave like AKEH in that AKEH can be used as a predicate whose subject is a nominalised predicate, but PIRANG-PIRANG is questionable, if not completely ill-formed.

(124)a. OLEHEKU MANGAN / AKEH.

*'I ate a lot.'* (lit. *my eating is much*)

b. ?OLEHKU MANGAN GEDANG / PIRANG-PIRANG.

*'I ate a lot of bananas.'* (lit. *my eating bananas is many*)

FORM 14: (KA)PING PIRA (berapa kali) - *'a number of times'*

The combination of (KA)PING (kali) and root is a productive process with numeral. It can also be applied to some of the derived forms above.

(125) KOWE / MANGAN (KA)PING PIRA?

Kamu / makan berapa kali?

(lit. *'You eat how many times?'*)

(126) KOWE / MANGAN (KA)PING PAPAT.

Kamu / makan empat kali.

*'You eat four times.'*

(127) KOWE / MANGAN (KA)PING PIRANG-PIRANG.

Kamu / makan beberapa kali. = Kamu / makan ber-kali-kali.

*'You eat many times.'*

The difference between AKEH and PIRANG-PIRANG shows up again when (KA)PING is used.

(128) \*KOWE / MANGAN (KA)PING AKEH.

to mean: *'You eat many times.'*

Indonesian allows an alternate form of beberapa kali in (127), namely, ber-kali-kali. This is a productive process. Compare:

(129)a. Saya / membeli beberapa cangkir kopi. =

Saya / membeli bercangkir-cangkir kopi.

*'I bought many cups of coffee.'*

b. Saya / berjalan beberapa mil. =

Saya / berjalan ber-mil-mil.

*'I walk many miles.'*

The form ber-NP-NP is form 13 above, so that it implies excessive amount. (129a) with ber-NP-NP is better translated as *'I bought cups (and cups) of coffee.'* Similarly, (129b), the second sentence is better translated as *'I walk miles (and miles).'* It is interesting to note that beberapa in Indonesian means *'some'* or *'indefinite amount'*, but when combined with an NP it means *'many'*, which is very similar to English *'some books'* although in English it does not seem to always

mean 'many books'. Furthermore, beberapa kali means 'many times', even though the notion 'indefinite number of times' is there, the emphasis is on 'numerous times'. How this differs from English 'indefinite some' and 'definite some', is not clear to me.

FORM 15: TIME-PIRA (TIME-berapa) - 'what-TIME'

The root can be combined with a word expressing time. However, PIRA (berapa) is only used when the expected answer includes a numeral. APA (apa) is used when the expected answer is a name such as 'Monday' or 'January'.

- (130) JAM } PIRA / SAIKI?  
TANGGAL }  
TAUN }
- Jam }  
Tanggal } berapa / sekarang?  
Tahun }
- 'What {  
time }  
date } is now/it?'  
year }

- (131) DINA }  
WULAN } APA / SAIKI?  
\*TAUN }
- Hari }  
Bulan } apa / sekarang?  
\*Tahun }
- 'What {  
day }  
month } is now/it?'  
year }

ROOT 4: KABEH (semua) - 'all'

Javanese	Indonesian	Gloss
1. KABEH	semua	'all'
2. KABEH-E	semua-nya	'all together' (?)
3. *KABEH-KABEH-E	*semua-semua-nya	*'at the (all)'
4. *KE-KABEH-EN	*ke-semua-an	*'too (all)'
5. *(SE-KABEH) *PADA KABEH-E	*se-semua	*'as (all) as'
6. *SAK-KABEH-KABEH-E	*se-semua-semua-nya	*'at the (all)'
7. -	*ber-semua-semua	*'in (all) groups'
8. KABEH-KABEH	*semua-semua	'all that one does'
9. -	*semua demi semua	*' (all) by (all)'
10. *ORA KABEH-KABEH-A	*tidak semua-pun	*'not the (all)'

As can be seen from the above diagram, root 4 is not as productive as the ones presented above.

FORM 1: KABEH (semua) - 'all'

The root form can only be used as a noun modifier. In the surface structure, the form may appear as something other than a noun modifier, but it is a result of a process recently known as Q-floating (see, for example, Keyser and Postal 1976).

(132) \*BOCAH KUWI / KABEH.

\*Anak itu / semua.

\*'The children are all.'

(133) KABEH BOCAH KUWI / AREP LUNGA.

Semua anak itu / akan pergi.

'All the children will go.'

(134) KABEH BOCAH / SENENG PERMEN.

Semua anak / suka gula-gula.

'All children like candy.'

Topicalisation can be applied to (133) and (134) to get (135) and (136), respectively:

(135) BOCAH KUWI // KABEH / AREP LUNGA.

Anak itu // semua / akan pergi.

'The children, (they) all will go.'

(136) BOCAH KUWI / KABEL / SENENG PERMEN.

Anak (itu) // semua / suka gula-gula.

'Children, (they) all like candy.'

FORM 2: KABEH-E (semua-nya) - 'all together'

The suffixed E (nya) form is syntactically a noun expressing a superordinate class, similar to 'to amount, the size', etc.

(137) BUKU KUWI // KABEH-E / PIRA?

Buku itu // semua-nya / berapa?

'The books, how much are they all together?'

Topicalisation seems obligatory, since the source (untopicalised) sentence seems ill-formed:

(138) \*KABEH-E BUKU KUWI / PIRA?

\*Semua-nya buku itu / berapa?

\*'How much are the all-together books?'



To reflect the noun-ness of the form, maybe a better translation is '*the total amount*' or '*the entire amount*'.

FORMS 3, 4, 5, 6, 7, 9, and 10:

These forms are not available since they are semantically impossible. The semantic constraints on these forms seem to be universal, but I do not know how to state them in a general way, at the time.

FORM 8: KABEH-KABEH (\*semua-semua) - '*all that one does (?)/anything*'

The reduplicated root, which is only available in Javanese, refers to a generic noun, act, etc. The closest Indonesian form is *apa saja* which corresponds more closely to Javanese *APA WAE* '*whatever*'.

- (139)  $\left. \begin{array}{l} \text{KABEH-KABEH} \\ \text{APA-APA} \\ \text{APA WAE} \end{array} \right\} / \text{ORA OLEH.}$   
 $\left. \begin{array}{l} *Semua-semua \\ ?Apa-apa \\ \text{Apa saja} \end{array} \right\} / \text{tidak boleh.}$   
 '*All that one does is not allowed.*'

- (140)  $\left. \begin{array}{l} \text{KABEH-KABEH} \\ \text{APA-APA} \\ \text{APA WAE} \end{array} \right\} / \text{DIPANGAN.}$   
 $\left. \begin{array}{l} *Semua-semua \\ ?Apa-apa \\ \text{Apa saja} \\ \text{Apa-pun} \end{array} \right\} / \text{dimakan.}$   
 '*All/anything/every single thing(s) were eaten.*'

ROOT 5: SABEN (setiap; masing-masing) - '*every; each*'

Javanese	Indonesian	Gloss
1. SABEN	setiap	'every'
2. SABEN-E	(biasanya)	'usually'
3. SABEN-SABEN	(kadang-kadang)	'sometimes'
4. (SIJI-SIJI-NE)	masing-masing	'each'
5. KE-SABEN-EN	(terlalu sering)	'too often'

The semantically related concepts '*every*', '*sometimes*', etc. are related in their forms in Javanese, but not in Indonesian. As is the case with root 4 (KABEH), the root is not as productive as the earlier roots.

## FORM 1: SABEN (setiap) - 'every'

The Javanese SABEN is used both for 'every' and 'each', but Indonesian has two separate words, setiap and masing-masing. When explicit distinction is needed, Javanese may use SIJI-SIJINE to express 'each' (see form 4).

(141) SABEN BOCAH / AREP LUNGA.

Setiap anak / akan pergi.

*'Every child will go.'*

(142) BOCAH KUWI / TURU SABEN AWAN.

Anak itu / tidur setiap siang.

*'The child sleeps every afternoon.'*

Even though plurality need not be overtly expressed, both languages use complete reduplication to overtly express plurality. If plural is overtly expressed, then it cannot be used with SABEN (setiap).

(143) \*SABEN BOCAH-BOCAH / AREP LUNGA.

\*Setiap anak-anak / akan pergi.

\**'Every children will go.'*

The above constraint applies to other Qs:

(144)

\*KOWE / DUWE BUKU-BUKU

{ AKEH.  
SETHITHIK.  
PIRA?  
PIRANG-PIRANG.  
LIMA.

\*Kamu / punya buku-buku

{ banyak.  
sedikit.  
berapa?  
(banyak sekali).  
lima.

'You have {  $\begin{matrix} \text{many} \\ \text{a few} \\ \text{how many} \\ \text{numerous} \\ \text{five} \end{matrix}$  } books.'

The ill-formedness of (143) is the result of a contradictory feature in SABEN (setiap) and plural noun, where SABEN (setiap) refers to singular item. The ill-formedness of (144) is the result of redundancy. These languages do not allow plural markers when a quantifier expressing plurality is already available.

## FORM 2: SABEN-E ((biasanya)) - 'usually'

The Javanese SABEN-E is related to SABEN in a manner that 'usually' can be interpreted as 'in every previous occasion'. The Indonesian

form *biasanya* comes from *biasa* meaning 'ordinary'. Sometimes the Javanese form is used as a compound form with *ADAT* meaning 'tradition'. As the meaning indicates, the form is used as an adverb of frequency or manner.

- (145) TANDANGANA KAYA (ADAT) SABEN-E.  
 Kerjakanlah seperti biasanya.  
*'Do it as usual/as is usually done.'*

- (146) SABEN-E / PIRA?  
 Biasanya / berapa?  
*'Usually, how much?'*

FORM 3: SABEN-SABEN ((kadang-kadang)) - 'sometimes'

The Javanese SABEN-SABEN is related to SABEN in a manner that 'sometimes' can be interpreted as 'every so often'. The Indonesian kadang-kadang is a root form, it does not come from the non-existing \*kadang. The form is an adverb of frequency.

- (147) BOCAH KUWI / SABEN-SABEN TURU AWAN.  
 Anak itu / kadang-kadang tidur siang.  
*'The child sometimes sleeps in the afternoon.'*

The Javanese SABEN-SABEN and the Indonesian kadang-kadang are not exactly the same, since Indonesian also makes use of *sering* meaning 'often'. It seems that SABEN-SABEN can be used for both or the meaning is probably in between. Javanese has *KEREP* meaning 'often'. The diagram below may reflect the meaning of SABEN-SABEN more precisely.

- (148) (never) < (sometimes) (often) > (always)  
 - < SABEN-SABEN KEREP > -  
 - < kadang-kadang sering > -

The use of SABEN-SABEN in sentences such as (149) seems neutral with regard to 'often' or 'sometimes':

- (149) BOCAH KUWI, SABEN-SABEN TILIKANA.  
*'That child, once in a while (?), visit him.'*

FORM 4: (SIJI-SIJI-NE) (masing-masing) - 'each'

The Javanese word to express 'each' is derived from a numeral *SIJI* 'one' and will be discussed later, since it belongs to Definite

Quantifiers. As stated above, Javanese does not necessarily distinguish 'each' and 'every'.

(150) Masing-masing anak / mendapat satu rupiah.

*'Each child gets one rupiah.'*

The form masing-masing is a root form, it does not come from the word \*masing. As is the case with setiap, overt pluralisation is not allowed, meaning that setiap and masing-masing can only modify a singular noun.

(151) \*Masing-masing anak-anak itu / mendapat satu rupiah.

*\*'Each children gets one rupiah.'*

The form can also 'float'. See Q-floating in Part II.

FORM 5: KE-SABEN-EN ((terlalu sering)) - 'too often'

The form KE-Q-EN in Javanese follows a general form, where KE--EN affixation carries the meaning of 'too-Q'. Recall the comparison of meanings in (148) above. With KE-SABEN-EN, the meaning is closer to KEREP. As a matter of fact, some speakers will use KE-KEREP-EN instead of KE-SABEN-EN.

(152) BOCAH KUWI / KE-KEREP-EN TILIK KANCANE.  
SABEN

Anak itu / terlalu sering menengok temannya.

*'The child visits his friends too often.'*

As is the case with most adverbs in Javanese, the frequency adverb KE-KEREP-EN or KE-SABEN-EN is used more often with nominalised predicate constructions such as:

(153) BOCAH KUWI // OLEHE TILIK KANCANE / KE-KEREP-EN.  
SABEN

*'The child, his visiting (of?) his friends is too often.'*

### B. DEFINITE QUANTIFIERS

Definite quantifiers refer to quantifiers that contain numerals. In terms of derivational productivity and semantic consequence of derivational processes, I will distinguish two roots: Root 1 is SIJI ('satu') 'one', and Root 2 is any numeral other than 'one'.

Since the structure of numerals is described in any language textbook or grammar, I will not describe it here.

## ROOT 1: SIJI (satu) - 'one'

Javanese	Indonesian	Gloss
1. SIJI	satu	'one'
2. SIJI-SIJI	satu (per/demi) satu	'one by one'
3. SIJI-SIJI (=/nasal/-SIJI)	?satu-satu	'one each'
4. SAK-WIJI-NE	sesuatu	'indefinite one/some'
5. (SING) SIJI-NE	yang satu (?nya)	'the other one'
6. SIJI-SIJI-NE	satu-satu-nya	'the only one'
7. SIJI-SIJI-NE	masing-masing	'each (one)'
8. ((KA)PING SIJI) = SEPISAN	satu } kali se	'one time (= once)'
9. -	ke-satu	'the first'
10. -	ke-satu-an	'a unit'
11. SALAH SIJI	salah satu	'one of them'
12. -	ber-satu	'united'
13. IJI-AN	(sendiri)	'alone/by oneself'

## FORM 1: SIJI (satu) - 'one'

The root SIJI (satu) is used as a noun quantifier or a predicate which only takes a quantity noun as its subject. The initial S (s)- may be a prefix if we compare with forms such as IJEN, SAKWIJINE, NGIJEN (instead of NYIJEN) in Javanese, or forms such as sesuatu and piatu in Indonesian. I am not familiar with their historical development.

- (154) AKU / DUWE BUKU SIJI.  
 Saya / punya buku satu.  
 'I have one book.'

Indonesian allows the reverse order in some dialects, but not Javanese.

- (155) Saya / punya satu buku.  
 \*AKU / DUWE SIJI BUKU.

The quantifier can be used as a predicate in Javanese by predicate nominalisation:

- (156) OLEHKU TUKU BUKU / SIJI.  
 'My buying books is one.'

In both languages, the form can be used as a predicate with countable noun (plural) as subject.

(157) BUKUKU / (MUNG) SIJI.

Buku saya / (hanya) satu.

*'My book is (only) one.'*

The use of MUNG (hanya) in (157) will make the sentence semantically more appropriate. Sentences such as (158) is grammatical but semantically odd (as odd as the English sentence seems to be):

(158) BUKU KUWI / SIJI.

Buku itu / satu.

*'The book is one.'*

The word SIJI, and probably also satu, means 'a piece' or 'one item' when used in constructions such as:

(159) BUKU KUWI // SIJI / LIMANG RUPIAH.

Buku itu // ?satu / lima rupiah.

*'The books, one is five rupiahs.'*

Even though (159) is clearly a topicalised construction, its source is not clear. Following a normal topicalisation rule, it should come from (160), but it has a different meaning from (159).

(160) BUKU SIJI KUWI / LIMANG RUPIAH.

?Buku satu itu / lima rupiah.

*'That one book is five rupiahs.'*

Another possible source is a generic sentence:

(161) BUKU SIJI / LIMANG RUPIAH.

?Buku satu / lima rupiah.

*'A book is five rupiahs.'*

FORM 2: SIJI-SIJI (satu-per/demi-satu) - 'one by one'

The reduplicated root is used primarily as a manner adverb.

(162) MLEBUA SIJI-SIJI.

Masuklah satu demi satu.

*'Come in one at a time/one by one.'*

In Indonesian, in addition to satu-demi-satu, some speakers use satu-per-satu or satu-satu. The morpheme per may be a borrowing, although it appears in words such as per-tama and se-per-tiga meaning '(the) first' and 'one third' respectively.

The reduplication can also signify emphatic expressions such as (Indonesian uses pun):

- (163) SIJI-SIJI LUMAYAN.  
 Satu-pun bolehlah.  
*'(Even if only) one, it's not bad.'*

FORM 3: SIJI-SIJI (= /nasal/-SIJI) (?satu-satu) - 'one each'

The reduplicated form to express 'one each' is used in sentences such as:

- (164) WENEHANA SIJI-SIJI.  
 ?Berilah satu-satu.  
*'Give (them) one each.'*

Javanese has NYIJI as an alternate form and Indonesian tends to use *satu seorang* 'one a person' which is a more productive form comparable to *satu-serumah* 'one in each house', etc.

- (165) GEDHANG KUWI // DUMEN SIJI-SIJI.  
 Pisang itu // berilah satu-seorang.  
*'The bananas, distribute them one each person.'*

FORM 4: SAK-WIJI-NE (sesuatu = suatu) - 'indefinite one/some'

Indonesian has more pro-forms (including pronouns) than Javanese. *Sesuatu* can be used as an independent pro-form but not SAK-WIJI-NE (= SAK-WIJI-NING).

- (166) Saya akan membeli sesuatu.  
 \*AKU AREP TUKU SAKWIJINE.  
*'I will buy something.'*

When used as a modifier of a noun, *suatu* is also allowed in Indonesian.

- (167) Saya melihat (se)suatu desa kecil dipuncak gunung itu.  
 AKU WERUH SAKWIJINING DESA CILIK ANA PUCUKING GUNUNG KUWI.  
*'I saw a small village on top of the mountain.'*

When used to refer to an indefinite place, time, etc. Javanese seems to allow more freedom in deleting the preposition than Indonesian.

- (168)a. ING SAKWIJINING DINA, BAPAKE DIPATENI.  
Pada suatu hari, ayahnya dibunuh.  
 b. SAKWIJINING DINA, BAPAKE DIPATENI.  
 ?Suatu hari, ayahnya dibunuh.  
*'One day, his father was murdered.'*

Indonesian also tends to use classifiers as pro-forms. See the discussion on Classifiers below.

FORM 5: (SING) SIJI-NE (yang satu (nya)) - *'the other one'*

This form is a relative clause form SING (yang). The availability of the suffix in Javanese makes it possible to delete SING, but the deletion of yang in Indonesian will make the phrase ambiguous, so yang tends to be retained.

(169) JUKUKNA (SING) SIJINE.

Ambilkan yang satu.

*'Get (me) the other one.'*

(170)a. JUKUKNA SIJINE.

*'Get (me) the other one.'*

b. JUKUKNA SIJI.

*'Get (me) one.'*

c. Ambilkan satu.

*'Get (me) one.'*

(171)a. (SING) SIJINE / AREP DIGAWA SESUK.

Yang satu / akan dibawa besok.

*'The other one will be brought tomorrow.'*

b. \*Satu / akan dibawa besok.

The above SING (yang) phrases are reduced relative clauses whose understood main nouns are deleted. The deletion is optional and the main noun (or head noun) may be overtly expressed:

(172) KURSI SING SIJINE / AREP DIGAWA SESUK.

Kursi yang satu / akan dibawa besok.

*'The other chair will be brought/taken tomorrow.'*

When the head noun is overtly expressed, it seems that in Javanese the deletion of SING makes the sentence less grammatical, and in my dialect of Indonesian the sentence is better if *nya* is added to either after *kursi* or *satu*.

(173)a. ?KURSI SIJINE / AREP DIGAWA SESUK.

b. Kursinya jang satu / akan dibawa besok.

c. Kursi yang satunya / akan dibawa besok.

d. \*Kursinya yang satunya / akan dibawa besok.

I was told by an Indonesian speaker that the need for *nya* is influenced by the corresponding requirement for *E* in Javanese. I will let dialectologists sort them out. If *nya* is part of my dialect, then a synchronic description of the language must consider it as an integral part of the system.



FORM 6: SIJI-SIJI-NE (satu-satu-nya) - *'the only one'*

The reduplicated root plus suffix E (nya) is used as a noun-modifier. When it appears in the surface structure as a predicate, it implies an understood noun being referred to.

- (174) ALI KUWI / SIJI-SIJI-NE BOCAH SING KENA DIPERCAYA.  
 Ali itu / satu-satu-nya anak yang dapat dipercaya.  
*'That Ali is the only child who can be trusted.'*

Javanese also uses SIJI-SIJI-NING instead of SIJI-SIJI-NE.

- (175) ?APA IKI / SIJI-SIJI-NE?  
 Apa ini / satu-satu-nya?  
*'Is this the only one?'*

Sentence (171) contains an understood NP. The Indonesian sentence seems more acceptable than the Javanese one.

FORM 7: SIJI-SIJI-NE (masing-masing) - *'each (one)'*

As stated above, Indonesian makes a covert distinction between *'each'* and *'every'*, but not Javanese (see Root 5, form 4). As usual, this situation invites borrowing, so that the use of SIJI-SIJI-NE to express masing-masing (*'each'*) is diminishing.

- (176) Masing-masing anak / mendapat satu rupiah.  
 ?SIJI-SIJI-NE BOCAH / OLEH SAKRUPIAH.  
*'Each child receives one rupiah.'*

Javanese sentence in (176) seems better with topicalisation:

- (177) Anak itu // masing-masing / mendapat satu rupiah.  
 BOCAH KUWI // SIJI-SIJINE / OLEH SERUPIAH.  
*'The children, each receives one rupiah.'*

FORM 8: ((KA)PING SIJI) = SEPISAN (satu kali = se-kali) - *'once'*

The morpheme (KA)PING (kali) is used productively to express multiplication. Like most languages these languages have special forms for one-times.

- (178) THUTHUKEN (KA)PING SIJI / SEPISAN.  
 Pukullah ?satu kali / sekali.  
*'Hit (it) once.'*

FORM 9: (ke-satu) - *'the first in a sequence/order'*

Only Indonesian has the generic term ke-satu *'the first'*, ke-dua *'the second'*, etc. Javanese only has a borrowed form NOMER SIJI, NOMER LORO, etc. These borrowed forms are sometimes translated back into Javanese into ANGKA SIJI, ANGKA LORO, etc. where ANGKA = nomer = *'number'*. The Indonesian ke-satu has an alternate form that is more frequently used. The form is pertama.

- (179)a. Ali itu / anak yang pertama.  
 b. Ali itu / anak yang ke-satu.  
 c. ALI KUWI / BOCAH (SING) MBAREP.  
*'Ali is the first child.'*

The Javanese MBAREP is only used in kinship systems, but pertama is used as a general form. There is also a derived form of pertama, namely pertama-tama *'first of all'*.

- (180) Pertama-tama, rumah ini harus dibersihkan.  
 ?SING PALING UTAMA, OMAH IKI KUDU DIRESIKI.  
*'First of all, this house must be cleaned.'*

The word pertama may come from utama *'the utmost'*.

FORM 10: (ke-satu-an) - *'a unit'*

The Indonesian affix ke--an is a productive affix. It is one of the nominal affixes in Indonesian. Compare:

- (181)a. satu → ke-satu-an (*'one' → 'a unit'*)  
 b. lahir → ke-lahir-an (*'born' → 'birth'*)  
 c. besar → ke-besar-an (*'big' → 'magnitude'*)

FORM 11: SALAH SIJI (salah satu) - *'one of them'*

Semantically SALAH (salah) *'numeral'* is similar to relative clauses in that it restricts the referent to a small number of objects. As a matter of fact with SALAH (salah) only SIJI (satu) can be used. Like relative clauses, the form is used as a noun-modifier, even though at the surface structure the understood head-noun may be deleted.

- (182) SALAH SIJI-NE BOCAH KUWI / KUDU LUNGA.  
 Salah satu anak itu / harus pergi.  
*'One of the children must go.'*

- (183) UNDANGNA SALAH SIJI.  
 Panggilkan salah satu.  
*'Call (for me) one of them.'*

## FORM 12: (ber-satu) - 'united'

Ber+NUMERAL is a productive process to express a unit indicated by the numeral. Ber-satu 'united', ber-dua 'in two', etc. With satu, the derived form has acquired a unique (although related) from the rest of ber+NUMERAL, so as a result, it uses another word sendiri ?'in one/alone', which corresponds to the Javanese form 13 below.

- (184) Bersatu kita teguh.  
'United we are strong.'

## FORM 13: IJI-N (sendiri) - 'alone/by oneself'

The N or AN suffix in Javanese is a productive suffix corresponding to Indonesian ber when used with numerals other than satu 'one'.

- (185) BOCAH KUWI / LUNGA IJEN.  
Anak itu / pergi sendiri.  
'The child went alone.'

- (186) BOCAH KUWI / IJEN.  
Anak itu / sendiri.  
'The child is alone.'

The form IJEN (sendiri) is used either as an adverb or an adjective, as the examples in (185) and (186) show.

Javanese also has an alternate form to express 'alone':

- (187) BOCAH KUWI / LUNGA DEWEKAN.  
'The child went alone.'

The word DEWEKAN comes from DEWE 'self', which is similar to Indonesian sendiri which comes from diri 'self'.

## ROOT 2: TELU (tiga) - 'three' &lt;or any number other than one&gt;

To contrast root 2 with root 1, the paradigm for root 1 is used below.

Javanese	Indonesian	Gloss
1. TELU	tiga	'three'
2. TELU-TELU	tiga-demi-tiga	'three by three'
3. /nasal/-TELU	*tiga-tiga	'for every three'
4. *SAK-TELU-NE	*se-tiga	'any three'
5. a. *(SING)-TELUNE	*yang tiganya	'the other three'
b. SING TELU	yang tiga	'the three'

	Javanese	Indonesian	Gloss
6.	a. *TELU-TELUNE	*tiga-tiganya	'the only three'
	b. TELU-TELUNE	ke-tiga-tiga-nya	'all three'
7.	-	-	'each three'
8.	(KA)PING TELU	tiga kali	'three times'
9.	ANGKA TELU	ke-tiga	'the third'
10.	-	-	'tripartite'
11.	a. *SALAH TELU	*salah tiga	'three of them'
	b. SING TELU	yang tiga	'the three' (see 5b)
12.	-	*ber-tiga	'union of three'
13.	TELU-AN (= TELON)	ber-tiga	'in three'

FORM 1: TELU (tiga) - 'three'

This form is used in exactly the same manner as SISI (satu). With regard to the overt plural form (reduplication) SISI (satu) cannot go with reduplicated nouns indicating plural because of semantic contradiction, and root 2 cannot occur with reduplicated nouns indicating plural because of semantic redundancy.

Root 2, but not root 1, can be combined with a noun to form a compound noun referring to measurement, duration, etc.

(188) AKU / AREP LUNGA TELUNG DINA.

Saya / akan pergi tiga hari.

'I will be away for three days.'

Note that in Javanese, a nasal linking sound must be added, and the Indonesian phrase tiga hari is not the same class as tiga buku 'three books'. The latter is an alternate form of buku tiga but the former is not an alternate form of \*hari tiga. Only tiga hari is available to express 'three days'. When root 1 is used, the alternate forms SE (se) must be used, instead of the root SISI (satu).

(189)a. \*AKU / AREP LUNGA SIJING DINA.

\*Saya / akan pergi satu hari.

b. AKU / AREP LUNGA SEDINA.

Saya / akan pergi sehari.

'I will be away for one day.'

Javanese also has a special (reduced) form for 'two' in compound constructions.

- (190) \*AKU / AREP LUNGA LORONG DINA. → AKU / AREP LUNGA RONG DINA.  
*'I will be away for two days.'*

FORM 2: TELU-TELU (tiga-demi-tiga) - *'three by three/three at a time'*

This form seems to behave like root 2. In constructions such as (163), however, it seems that root 1 is better than root 2, maybe for a semantic reason. The phrase *'even if only ...'* is progressively worse when continued with larger numerals. Compare the following:

- (191)a. SIJI-SIJI / LUMAYAN.  
 Satu-pun / bolehlah.  
*'(Even if only) one will do.'*
- b. ?TELU-TELU / LUMAYAN.  
 ?Tiga-pun / baiklah.  
*'(Even if only) three will do.'*

FORM 3: ?/nasal/-TELU (\*tiga-tiga) - *'for every three/three each'*

I am not sure whether this form is allowed in Javanese, since I do not use it at all. The form is plausible. I can accept NYIJI but not NELU.

- (192)a. GEDHANG KUWI DUMEN NYIJI.  
*'Distribute the bananas, one for each (person).'*
- b. \*GEDHANG KUWI DUMEN NELU.  
*'Distribute the bananas, three for each (person).'*

FORM 4: \*SAK-TELU-NE (\*se-tiga) - *'any three'*

In English *'any three'* means an indefinite or unspecified object of definite amount (i.e. three). Both Indonesian and Javanese do not have a similar expression. As noted above, Indonesian has *sesuatu* *'something/anything'*, but the formation is not productive. There is a general term for *'anything'*: SAK-SAK-E (apa saja) *'whatever'*. To express *'any three'*, for example, Javanese and Indonesian can get around it this way:

- (193)a. AKU BUTUH BUKU TELU.  
*'I need three books.'*
- b. SING ENDI?  
*'Which ones?'*
- c. SAK-SAK-E.  
*'Whichever.'*

- FORM 5: a. \*(SING)-TELU-NE (\*yang tiga-nya) - 'the other three'  
 b. SING TELU (yang tiga) - 'the three'

Although 'the other three' is logically possible, neither language has an expression formed in the same manner as the one to express 'the other one', namely form 5a: (SING)-Q-E (yang-Q).

- (194) \*JUKUKNA (SING) TELUNE.  
 \*Ambilkan yang tiganya.  
 'Get (me) the other three.'

Form 5b is available. It is a regular relative clause construction.

- (195) BUKU SING TELU KUWI / AREP TAK DOL.  
 Buku yang tiga itu / akan saya jual.  
 'Those three books will be sold by me. = I will sell those three books.' (lit. The books which are three will be sold by me)
- (196) BUKU TELU KUWI / AREP TAK DOL.  
 Buku tiga itu / akan saya jual.  
 'I will sell those three books.'
- (197) BUKU KUWI // SING TELU / AREP TAK DOL.  
 Buku itu // yang tiga / akan saya jual.  
 'I will sell three of my books.' (lit. The books, three of them will be sold by me)

When a complete relative clause is used, it can be restrictive or non-restrictive, as in (195). The English gloss is misleading because (195) can mean either that there are more than three books or there are only three books and the speaker will sell them all. When the clause is reduced, as in (196), then the ambiguity seems to disappear. BUKU TELU KUWI (Buku tiga itu) implies that there are only three books that the hearer can refer to, i.e. visible to him/her. When topicalised constructions are used, as in (197), the clause is definitely restrictive.

- FORM 6: a. \*TELU-TELUNE (\*tiga-tiganya) - 'the only three'  
 b. TELU-TELUNE (ke-tiga-tiganya) - 'all three'

Form 6a is restricted to root 1. In Javanese, when used with root 2 the only meaning available is 'all NUMERAL' which has the corresponding Indonesian form ke-reduplicated root.

(198) \*BOCAH-BOCAH KUWI / TELU-TELUNE BOCAH SING KENA DIPERCAYA.

\*Anak-anak itu / tiga-tiganya anak yang dapat dipercaya.

*'Those children are the only three children who can be trusted.'*

(199) BOCAH KUWI / TELU-TELUNE KENA DIPERCAYA.

Anak itu / ke-tiga-tiga-nya dapat dipercaya.

*'Those children, all three can be trusted.'*

The Indonesian ke-Q-Q-nya cannot be used with root 1.

(200) \*Anak itu / ke-satu-satu-nya dapat dipercaya.

FORM 7: *'each three/every three'*

With root 1, Indonesian has masing-masing to express *'each one'*, but there is no expression for *'each three'*. I am not sure if setiap tiga can be used to express *'every three'*. Javanese and Indonesian have forms to express *'every-Q-MEASUREMENT'*:

(201) TANDURAN KUWI SIRAMANA SABEN TELUNG DINA.

Tanaman itu berilah air setiap tiga hari. (?)

*'Water the plant every three days.'*

FORM 8: (KA)PING TELU (tiga kali) - *'three times'*

The form (KA)PING-Q (Q-kali) can be used with any numeral.

FORM 9: ANGKA TELU (ke-tiga) - *'the third'*

This form is also used with any numeral.

FORM 10: *'tripartite'*

The form ke-Q-an in Indonesian can only be used with root 1. Similar words occur but not derived from the same process and mostly borrowings. For example: DWI-TUNGGAL *'a pair'* (2 but 1), sepasang *'a pair'*, etc.

FORM 11: a. \*SALAH TELU (\*salah tiga) - *'three of them'*

b. SING TELU (yang tiga) - *'the three (of them)'*

While *'any one of them'* is expressed by SALAH-Q (salah-Q), it only occurs with SIJI (satu) *'one'*.

(202) \*UNDANGNA SALAH TELU.

\*Panggilkanlah salah tiga.

*'Call (for me) any three (of them).'*

Form 11b is used to express '*the three of them*', etc. See form 5 above for a discussion of SING (yang)-Q.

FORM 12: (\*ber-tiga) - '*in three*'

Similar to ke-Q-an (form 10 above), the form ber-Q is only available with root 1, to express '*to be in -Q*'. The form is available, but has a different meaning.

FORM 13: TELU-AN (ber-tiga) - '*in three (just the three of them)*'

There is an interesting unparallelism between form 12 and form 13. The Javanese Q+AN with root 1 (and most of root 1 forms) corresponds to ber-Q in Indonesian, but not if Q is satu (root 1). Compare:

(203) Root 1:

Form 12: -                ber-satu                '*unit of one/united*'

Form 13: IJI-AN        (sendiri)                '*by oneself*'

Root 2:

Form 12: -                -                '*unit of three*'

Form 13: TELU-AN        ber-tiga                '*threesome*'

One possible explanation is that ber-satu was originally ambiguous but one meaning became so prominent that the second meaning is out of use, particularly with the help of the alternate form *sendiri*. Diachronic linguists can say more about this, I am sure. Although TELU-AN is the analogy of IJI-AN (IJEN), the former is used only in fractions such as SAK-PER-TELON '*one third*', while the Indonesian ber-Q can be used as an adverb.

(204) Mereka pergi bertiga.

\*WONG KUWI LUNGA TELON.

'*They went in threesome.*'

### C. CLASSIFIERS

Indonesian makes use of classifiers, in addition to quantifiers, to refer to the quantity of a noun. In formal writing, some grammarians consider classifiers mandatory, but in spoken language only a few are still in use. In general, a noun phrase has the structure (where CL = classifier):

(205) NP → Q+CL+N(Det)



For example:

dua-orang-anak(itu) 'two-CL-children-the'

dua-buah-batu(itu) 'two-CL-rock-the'

The following is a list of quantifiers in Indonesian with the approximate rule of usage for each, taken from Slametmuljana 1969:300-1.

Classifier	Types of nouns used	Examples
1. batang	something long	'tree, stick, cigarette'
2. bentuk	curvey objects	'ring, bracelet'
3. belah	one number of a pair	'eye, ear'
4. bidang	wide and flat objects	'land, sail'
5. biji	small objects	'cucumber, seed'
6. bilah	small, flat, sharp objects	'knife, sword'
7. buah	fruits and irregularly shaped objects	'town, island'
8. butir	small, round objects	'pearl, egg'
9. ekor	animals	'monkey, cow, bird'
10. helai	sheets of objects	'paper, cloth'
11. kuntum	objects with stems	'flower'
12. kaki	objects with handles	'umbrella'
13. keping	thin, solid objects	'tile, board, plank'
14. orang	people	'child, woman, man'
15. patah	part of language	'word, sentence'
16. {pintu tangga atap}	houses	
17. potong	portions of solid objects	'stick, wood'
18. penggal kerat	portions of broken objects	'stick, cloth'
19. utas	small, lean, long objects	'rope, hair'
20. pucuk	objects with pointed part (also: letter)	'gun, cannon, pen'
21. rawan	net	
22. rumpun	a group of plants	'banana trees'
23. tangkai	objects with stalks	'flower'
24. carik	portions of sheets of objects	'paper'

BIBLIOGRAPHY

AKMAJIAN, A. and A. LEHRER

- 1976 'NP-like Quantifiers and the Problem of Determining the Head of an NP'. *Linguistic Analysis* 2/4:295-414.

CHUNG, Sandra

- 1976 'An Object-creating Rule in Bahasa Indonesia'. *Linguistic Inquiry* 7/1:41-87.

IOUP, Georgette

- 1976 'Some Universals for Quantifier Scope'. In: J.P. Kimball, ed. *Syntax and Semantics*, vol.4:38-58. New York: Academic Press.

JACKENDOFF, Ray S.

- 1968 'Quantifiers in English'. *Foundations of Language* 4/4: 422-42.
- 1971 'On Some Questionable Arguments about Quantifiers and Negation'. *Language* 47/2:282-97.
- 1972 'Any vs. Every'. *Linguistic Inquiry* 3/1:119-20.

JOHNSON, D.E.

- 1974 Toward a Theory of Relationally-based Grammar. Ph.D. dissertation, University of Illinois at Urbana-Champaign.

KEYSER, S.J. and P.M. POSTAL

- 1976 *Beginning English Grammar*. New York: Harper and Row.

## LAKOFF, G.

- 1971 'On Generative Semantics'. In: D.D. Steinberg and L.A. Jakobovits, eds *Semantics*, 232-96. London: Cambridge University Press.

## LAWLER, J.M.

- 1975 'A Agrees with B in Achenese: A Problem for Relational Grammar'. (mimeo), University of Michigan. To appear in Peter Cole and J.M. Sadock, eds *Syntax and Semantics*, vol.8: *Grammatical Relations*. New York: Academic Press.

## SLAMETMULJANA

- 1969 *Kaidah Bahasa Indonesia*. Ende, Flores: Nusa Indah.

## SOEMARMO, M.

- 1970 Subject-Predicate, Topic-Comment, and Focus-Presupposition Constructions in Bahasa Indonesia and Javanese. Ph.D. dissertation, University of California at Los Angeles.
- 1975 'The Illusive Simple Nounphrase'. In: J.W.M. Verhaar, ed. *Miscellaneous Studies in Indonesian and Languages in Indonesia* 1:22-8. Jakarta: Badan Penyelenggara Seri NUSA.
- 1977 'Syntactic and Semantic Wellformedness'. In: J.W.M. Verhaar, ed. *Miscellaneous Studies in Indonesian and Languages in Indonesia* 3:19-26. Jakarta: Badan Penyelenggara Seri NUSA.



## VOWEL PATTERNING AND MEANING IN MALAY PAIR-WORDS

THAM SEONG CHEE

There is a class of words in Malay which Zainal Abidin Ahmad (Zaba) (1924) termed pair-words and which perhaps comes closest to phonetic symbolism as it is understood in the literature available on psycholinguistics. The term pair-words was defined as "set phrases consisting of two words combined which retain fully their literal meaning." Two types of pair-words were identified, namely (a) the alliterative, and (b) the non-alliterative. Those which come under the first category are pairs of words which repeat some of the sounds or reproduce them in some modified form as for example:

bukit} bukau	bukit-bukau ('hills and valleys')
gerak} geri	gerak-geri ('movements')
kueh} mueh	kueh-mueh ('cakes, savouries, puddings')

The first word is regarded as the principal word, and the second: qualifying, expanding, or intensifying the meaning of the first. What is meant is that when each pair-word is split, the morpho-syntactical status of the first word is primary whereas the second or following word is morpho-syntactically secondary. Words of primary status in this context have freedom of occurrence in other morpho-syntactical environments, while those words which are secondary do not have this freedom outside the fixed attachments they have with their respective primary or principal word. Thus the secondary word has semantic value, particularly when paired with the first, however, it cannot be treated in the conventional analytical fashion based on a contrast between signifier and signified. Zainal Abidin had suggested that formation

of secondary words largely depends on euphony whose emphasis is on agreement between sounds being motivated by such considerations as being pleasing to the ear and the ease of pronunciation.

In the case of pair-words which are non-alliterative, it is not a repetition of sounds or syllables by the second word that occurs but rather the role of the second word is essentially to signify a cognate idea on the basis of the first, e.g.:

kurus	'thin'	}	'thin and haggard'
kering	'dry'		
lintang	'across (of a barrier)'	}	'topsy-turvy, criss-cross, disorder'
pukang	'fork, junction'		
sangkut	'stick or stuck, correct'	}	'connection, implication'
paut	'clinging, hanging'		

Zaba explained non-alliterative pair-words as being arbitrarily structured and that they are all fixed conventional forms. The functions of such linguistic devices as found in alliterative and non-alliterative pair-words as suggested by him are (a) similar to those served by reduplication to strengthen and intensify the ordinary meaning of the primary single word, by adding to it an idea of indefinite plural confusion, (b) to imply indefinite repetition, association, continuity or multiplied quantity, or (c) when the pairs are nouns, to express an indefinite universal inclusion of all kinds or species (and sometimes all materials).

However, Zaba had tended to account for the existence of alliterative and non-alliterative pair-words in a general manner, attributing their phonological characteristics as being due to euphony. In this respect, his analysis is unsatisfactory and in the context of ascertaining phonetic symbolism in Malay inadequate. Pair-words in Malay, whether they are alliterative or non-alliterative exhibit a remarkable consistency of pattern and as such are semantically pertinent.

There are two ways of describing Malay pair-words. One way is to look at their syllabic structures and the other way is to examine their vowel and consonant contrasts. In the latter case some attempt will be made to link vowel contrast in Malay pair-words to phonetic symbolism.

Structurally, Malay pair-words may be regarded as a special kind of compound. Without exception two words or morphemes are involved in their formation. If only semantic criteria are used to determine compound status (such as when it is said that two words each with its own semantic value when linked gives a third semantic value), Malay pair-words would strictly be excluded as compounds. This is because the second word in the pair in Malay pair-words usually has no independent

semantic value. However, the second word is obligatory if the intention is to extend the semantic boundary of the first word. But the higher level of meaning signalled when two words are used as a pair retains substantially the central idea represented by the first word. We may regard compounds in Malay as made up of two types, viz. the 'true' compounds and the 'quasi' compounds. Malay pair-words come under the latter category. The use of the terms (true vs. quasi) is purely dictated by functional-explanatory reasons and does not suggest or imply a qualitative difference between the two. Also other dichotomies are possible outside the two suggested. Perhaps some examples to illustrate their difference would be useful at this juncture.

#### True Compounds

rumah	'house'	}	(rumah-tangga)	'a homestead'
tangga	'house-ladder, stairs; steps'			
meja	'table'	}	(meja-kerusi)	'furniture'
kerusi	'chair'			
kaki	'leg'	}	(kaki-tangan)	'staff; workers in an office'
tangan	'hand'			

#### Quasi Compounds (Pair-words)

anak	'child'	}	(anak-pinak)	'descendants; children of both sexes'
pinak	(no independent semantic value)			
senang	'ease; easy'	}	(senang-lenang)	'at ease; care-free; free of worries'
lenang	(no independent semantic value)			
batu	'stone; rock; pebble'	}	(batu-batan)	'all varieties and shades of stone'
batan	(no independent semantic value)			

#### Syllable Contrast in Pair-words

Syllable contrast in Malay pair-words may be best treated as linguistic phenomena associated with rhyming and chiming. In the case of rhyming a syllable in the primary word (which is always in initial position) matches a syllable in the secondary word (always in final position). Rhyming can occur between the initial syllable of the primary word and the initial syllable of the secondary word or between the final syllable of the primary word and the final syllable of the secondary word in disyllabic forms. In rare instances, rhyming may occur between the middle syllable of the primary word and the middle syllable of the secondary word in tri-syllabic forms, but this usually also involves the simultaneous rhyming of either the first or final syllable of the primary and secondary words as in:

se(rem)bah-se(rem)beh 'weeping copiously'  
se(ka)li-se(ka)la 'once in a while, occasionally'

The following are some examples of rhyming found in disyllabic forms:

#### Initial Rhyming

##### (a) Nouns

batu-batan 'all varieties and shades of stones'  
bukit-bukau 'hills and valleys'

#### Final Rhyming

##### (b) Nouns

kuih-muih 'all varieties of cakes'  
selok-belok 'intricacies of a problem'

#### Initial Rhyming

##### (a) Adjectives

gelap-gelita 'total darkness'  
jenggal-jenggul 'bumpy of roads'

#### Final Rhyming

##### (b) Adjectives

kaya-raya 'wealthy'  
chondong-mondong 'inclining precipitously'

#### Initial Rhyming

##### (a) Verbs

beli-belah 'shopping; buying'  
lesap-lesup 'vanished without a trace'

#### Final Rhyming

##### (b) Verbs

cerai-berai 'separated; scattered'  
halau-balau 'drive away furiously'

It should be noted that pair-words whatever their grammatical class may not be used as base forms to form complex words. In addition to this, pair-words are largely used (a) to show variety or heterogeneity, in which case they are verbs or nouns, and (b) to intensify or emphasise a quality or action in which case they are adjectives or adverbs. There are a great number of pair-words which are in fact onomatopoeic in function and character.

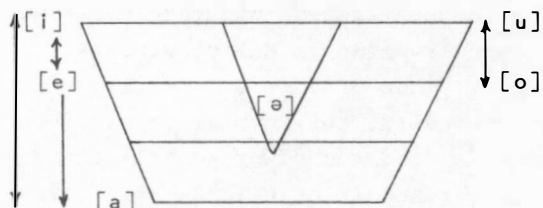
Chiming is the process whereby vowels in the primary word are somehow similar and these contrast with vowels in the secondary word which are also similar, e.g.:



asal-usul 'origins; beginnings'  
suku-sakat 'clans; relatives and clans'  
cucu-cicit 'grandchildren (of both sexes)'

While rhyming and chiming are useful concepts in describing Malay pair-words, their usefulness is however limited. This is because a great number of forms or pair-words cannot be accounted for by the use of these two concepts. To account for pair-words satisfactorily, vowel patterns underlying Malay pair-words would have to be described. The remaining section of this study will therefore attempt to show some coherence of the vowel-patterns as found in Malay pair-words. The concept vowel-patterning envisages the occurrence of certain vowel combinations in Malay pair-words, that is to say, the tendency for vowels in the secondary word to be somehow determined by the vowels in the primary word. Thus vowels in the primary word and vowels in the secondary word contrast in predictable ways. Contrast may also involve vowels and diphthongs. In certain instances all the vowels in both the primary and secondary words may also be similar. The patterns of contrast are basically structured in the manner shown by the vowel charts below:

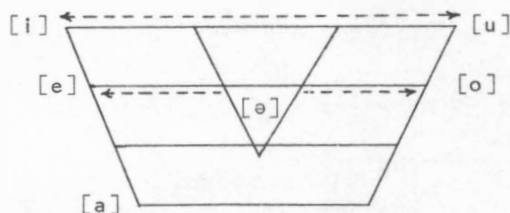
(a) Vertical Contrast or Patterning



[i] + [a] = [ia] becomes a vowel cluster

[a] + [i] = [ai] becomes a diphthong

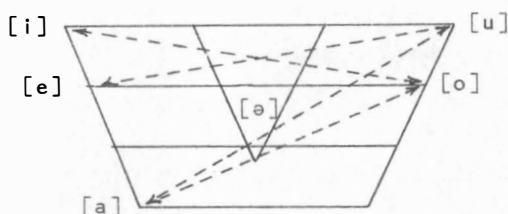
(b) Horizontal Contrast or Patterning



[i] + [u] = [iu] becomes a vowel cluster

[u] + [i] = [ui] becomes a diphthong

## (c) Cross-wise Contrast or Patterning



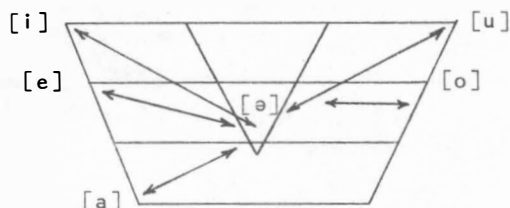
[i] + [o] = [io] becomes a vowel cluster

[o] + [i] = [oi] becomes a diphthong

[u] + [a] = [ua] becomes a vowel cluster

[a] + [u] = [au] becomes a diphthong

## (d) Contrast of Front and Back Vowels with the Neutral (Schwa) Vowel



We shall now look at two sets of phonological contrast employed in alliterative and non-alliterative pair-words in Malay, viz. (a) contrast made of consonants, and (b) contrast made of vowels. In almost every instance of contrast involving consonants, the initial sound of the primary word and the initial sound of the secondary word is the same sound. This characteristic is also found in pair-words which are constituted of two words having equal morphological status, that is to say having the freedom to occur as free forms in other morpho-syntactical environments such as for example:

tebang	'to fell large trees'
tebas	'to cut down scrub and bushes'
tebang-tebas	'to cut down trees, large and small'
tegap	'sturdy'
tegun	'stand firmly (especially when surprised or frightened)'
tegap-tegun	'well-built of person'

Thus there is a co-occurrence of an unvoiced plosive (alveolar, bilabial or velar) with an unvoiced plosive or alternately a voiced plosive with a voiced plosive. Similarly there is a co-occurrence of a fricative

(labio-dental, dental, alveolar, palato-alveolar, glottal) with a fricative of similar articulatory characteristics - voiced or unvoiced as the case may be. This principle of phonological distribution applies to laterals and affricates as well. Such a patterning of consonants is however not perfectly consistent as different consonants may co-occur in certain pair-words. It is however the vowels, both in terms of their patterning and opposition that provide the clue to phonetic symbolism in Malay. A number of patterns are observable and they may be systematised in the following order:

(a) Contrast of vowels occurs in the second or final syllable but the vowel in the first or initial syllable (in forms which are disyllabic) is similar, e.g.:

gunong-ganang ( <i>'mountains and valleys'</i> )	([o] contrasts with [a])
teka-teki ( <i>'riddles'</i> )	([a] contrasts with [i])
lalu-lalang ( <i>'pass to and fro'</i> )	([u] contrasts with [a])
bengkang-bengkok ( <i>'meandering'</i> )	([a] contrasts with [o])
celum-celam ( <i>'tramping in and out'</i> )	([u] contrasts with [a])

This form of contrast is the most popular and thus numerous examples may be obtained.

(b) Contrast of vowels occurs in both the first and second syllables of disyllabic forms showing perfect phonological balance, e.g.:

huru-hara ( <i>'tumult; uproar'</i> )	([u] contrasts with [a])
duduk-dadak ( <i>'sitting precariously'</i> )	([u] contrasts with [a])
hulur-halar ( <i>'extending; slacking'</i> )	([u] contrasts with [a])
hukum-hakam ( <i>'all varieties of law'</i> )	([u] contrasts with [a])
ongkoh-angkah ( <i>'tottering'</i> )	([o] contrasts with [a])
porok-parak ( <i>'running and stumbling'</i> )	([o] contrasts with [a])
susup-sasap ( <i>'stealthily'</i> )	([u] contrasts with [a])

gopoh-gapah ([o] contrasts with [a])  
(*'hurriedly'*)

huyung-hayang ([u] contrasts with [a])  
(*'tottering of gait'*)

(c) Vowel patterning showing perfect symmetry of form, e.g.:

senget-menget ([e] is consistent throughout)  
(*'a slant in an extreme way'*)

condong-mondong ([o] is consistent throughout)  
(*'sloping; inclining'*)

comot-momot ([o] is consistent throughout)  
(*'extreme filth; untidiness'*)

ceret-meret ([e] is consistent throughout)  
(*'like water, wateriness'*)

kaya-rayā ([a] is consistent throughout)  
(*'wealthy; opulent'*)

remeh-temeh ([e] is consistent throughout)  
(*'minor; trivial'*)

(d) Contrast of vowels occurs in the second or third syllable of tri-syllabic words but the vowels in the first and second syllables of first and third syllables are similar, e.g.:

ketunjang-ketunjit ([a] contrasts with [i])  
(*'clashing of sounds'*) in the third syllable

sakali-sakala ([i] contrasts with [a])  
(*'once in a while'*) in the third syllable

serebah-serebeh ([a] contrasts with [e])  
(*'hanging down, slovenly'*) in the third syllable

gelusor-gelasor ([u] contrasts with [a])  
(*'sliding; slipping'*) in the second syllable

jelepah-jelepoh ([a] contrasts with [o])  
(*'flop down'*) in the third syllable

celengkang-celengkak ([a] contrasts with [o])  
(*'zigzagging'*) in the third syllable

It will be noticed that vowel contrast in tri-syllabic forms occurs much more frequently in the third or final syllable. In any case alliterative or non-alliterative pair-words which are structurally tri-syllabic are relatively rare in contrast to those which are disyllabic - a linguistic fact noted earlier.

(e) Contrast of vowels or between vowel and diphthong occurs within the primary word and within the secondary word but together exemplifying perfect phonological balance in repetitive distribution, e.g.:

<u>hina-dina</u> ( <i>'lowly; common, of people'</i> )	([i] contrasts with [a])
<u>halai-balai</u> ( <i>'sixes and sevens; neglected'</i> )	([a] contrasts with [ai])
<u>halau-balau</u> ( <i>'drive away furiously'</i> )	([a] contrasts with [au])
<u>senang-lenang</u> ( <i>'at ease, extreme comfort'</i> )	([ə] contrasts with [a])
<u>sending-mending</u> ( <i>'heeling over'</i> )	([ə] contrasts with [i])
<u>sayor-mayor</u> ( <i>'all varieties of vegetables'</i> )	([a] contrasts with [o])

(f) Contrast of vowels may occur in the first syllable of disyllabic forms. This is however extremely rare as in the examples:

<u>haru-biru</u> ( <i>'uproar; tumult'</i> )	([a] contrasts with [i])
<u>jimat-cermat</u> ( <i>'frugality'</i> )	([i] contrasts with [ə])

(g) Finally contrast of vowels may occur in forms which have vowel clusters. In this case the first vowel in the first word contrasts with the first vowel in the second word and the second vowel in the first word contrasts with the second vowel of the second word, e.g.:

<u>cuak-caik</u> ( <i>'extreme timidity'</i> )	([u] contrasts with [a]) ([a] contrasts with [i])
<u>gediang-gediut</u> ( <i>'sound of band'</i> )	([i] contrasts with [i]) ([a] contrasts with [u])

We are now in a position to make certain statements about the use of vowels and at the same time gauge their distribution in order to get an understanding of their role in transmitting meaning - their phonetic symbolism. The examples utilised so far to illustrate the various patterns and characteristics of vowel distribution in Malay pair-words are by no means exhaustive, but they are typical. Therefore statements made while they will be derived from the examples cited will also implicitly account for examples not included.

The most significant feature that may be stated is that vowel contrast (in both alliterative and non-alliterative pair-words) involves front vowels (particularly [i], [e] and [a]) with back vowels (particularly [u] and [o]). The order is however not consistent throughout, in that a back vowel may precede a front vowel and vice-versa.

Secondly, contrast of vowels may occur between a high front vowel (for example [i]) and a low front vowel (for example [a]). Here again the order is not consistent as either vowel can occur first.

Thirdly, contrast of vowels may occur between a high back vowel (for example [u]) and a semi-high back vowel (for example [o]). Order is similarly characteristic in that either can occur first.

Fourthly, contrast of vowels may occur in a cross-wise manner between a low front vowel (for example [a]), and a high, back or semi-high, back vowel (for example [u] and [o]). Or alternately contrast is between a high front vowel [for example [i]] and a semi-high back vowel (for example [o]).

Fifthly, a similar vowel may occur throughout all the syllables of a pair-word as shown in section (c) or in the primary word or the secondary word respectively as shown in section (b).

Sixthly, contrast of vowels may occur between the high front vowel [i] or the semi-high front vowel [e] with the central or neutral vowel [ə] in both close or open syllables. In examples where the final syllable in pair-words is closed contrast between the low front vowel [a] and the neutral vowel [ə] also occurs. Contrast between the neutral vowel and the back vowels is relatively rare.

Seventhly, contrast between vowels and diphthongs may also occur. However, here too we find that the previous rules established (rules 1 and 2 in this discussion) exercise a consistent pressure on the patterns of contrast that may be permitted as for example [a] contrast with [ai] where [a] and [i] together constitute the diphthong. In another example [a] contrasts with [au] where [a] and [u] together constitute the diphthong. Jones (1960:58ff) defines a diphthong as "an independent vowel-glide not containing within itself either a 'peak' or a 'trough' of prominence". A diphthong is thus regarded as a syllable.

Finally, contrast of vowels in a cluster within the primary word; within the secondary word; and between the vowel cluster in the primary word and the vowel cluster in the secondary word exemplify the same degree of conformity with the patterns of vowel contrast found in rules 1 and 2 discussed earlier. Examples of contrast have been shown in the pair-words cuak-caik and gediang-gediut. A vowel cluster is different

from a diphthong in that it has as many peaks of prominence as there are vowels in the cluster.

We may conclude on the basis of the discussion thus far made that the vowel patterns (both in contrast or in conformity with each other) of Malay pair-words show a predictable sequence. Vowel patterning of the types that have been identified are thus harmonious. They exemplify in other words vowel harmony, a linguistic phenomenon Gleason (1961:84) defines as "the commonest type of non-contiguous assimilation, in which vowels of successive syllables must be similar in some way". Vowel harmony can be extensive or limited depending on the language. However, Gleason's description of vowel-harmony stressed the grammatical function of vowels in which semantic and grammatical differentiation runs parallel with vowel differentiation. In the case of Malay, the distribution of vowels and their patterning in pair-words do not exhibit the same degree of semantic and grammatical function. While vowel patterns as they are found in Malay pair-words are highly regular, yet vowels in terms of their distribution and patterning are not determined by rules of grammar but rather by preferred phonological contrast in the patterns as identified. Vowels in the patterns as they are found in Malay pair-words have a symbolic aspect suggesting meanings as identified by Zaba, enumerated at the beginning of this discussion. Meanings are therefore suggested or signalled by a subtle combination of sounds involving vowels, consonants and diphthongs in predictable sequences and patterns. The phenomenon of phonetic symbolism as it refers to pair-words in Malay may be understood in terms of the dichotomy focus and elaboration. The first or primary word (with its structure or pattern of sounds) draws attention to the intention, quality, situation or state-of-being suggested by it whereas the secondary word (with its structure and pattern of sounds closely cohering with those of the first) elaborates on it. Some examples are found in the following pair-words:

Focus	Focus + Elaboration
senang ( <i>'ease; easy'</i> )	senang-lenang ( <i>'at leisure; entirely comfortable'</i> )
sekali ( <i>'once'</i> )	sekali-sekala ( <i>'seldom, once in a while'</i> )
susup ( <i>'crawl under; pass under'</i> )	susup-sasap ( <i>'enter stealthily'</i> )
selang ( <i>'passage, intervening'</i> )	selang-seli ( <i>'alternately'</i> )

Focus	Focus + Elaboration
licin ( <i>'smooth; sleek'</i> )	licin-licau ( <i>'slippery; slipping away easily'</i> )
sungkor ( <i>'fall face forward'</i> )	sungkor-sangkar ( <i>'sprawling on one's face'</i> )

In the case of pair-words which are onomatopoeic, the vowel sounds are the key elements in symbolising meaning, although they are not structured of a focus and an elaboration in the real sense of the word. In such instances, there is a pseudo focus-elaboration relationship between the two words, e.g.:

lecup-lecap	<i>'sounds made by darts entering water'</i>
porok-parak	<i>'sounds of the feet slipping in a fight'</i>
lepah-lepoh	<i>'stumbling languidly'</i>
dentam-dentum	<i>'banging of guns or falling articles'</i>
dentang-dentong	<i>'booming of guns'</i>

Among pair-words which are non-onomatopoeic, the most perfect combination of words in a pair is that which is constituted of a focus and an elaboration; where the consonants are similar in form; and where there is only one feature of vowel contrast as in the examples:

Focus	Focus + Elaboration
sungkor ( <i>'fall forward'</i> )	sungkor-sangkor [sun̩kor-saŋkor] contrast is between [u] and [a] ( <i>'sprawling'</i> )
sekali ( <i>'once'</i> )	sekali-sekala [səkalɪ-səkələ] contrast is between [i] and [ə] ( <i>'rarely; occasionally'</i> )
dongkor ( <i>'bundle out'</i> )	dongkor-dangkar [doŋkor-daŋkar] contrast is between [o] and [a] ( <i>'expel neck and crop'</i> )

Individual phones or sounds, whether vowel, consonant or diphthong do not suggest or symbolise semantic values but it is the distribution of the sounds in pair-words which functions to expand, intensify or give a sense of continuous action to the meaning as contained in the primary word, namely, the focus.



BIBLIOGRAPHY

GLEASON, H.A., Jr

- 1961      *An Introduction to Descriptive Linguistics*. N.Y.: Holt,  
Rinehart & Winston. (First edition 1955.)

JONES, Daniel

- 1960      *An Outline of English Phonetics*. Ninth edition.  
Cambridge: Heffer.

ZAINAL ABIDIN AHMAD

- 1924      'Pair-words in Malay'. *Journal of the Malayan Branch of  
the Royal Asiatic Society* 5:324-38.



## EXPRESSIVES IN KEDAH MALAY

JAMES T. COLLINS

Through the years several books and articles have been written about Malay grammar. Some specifically deal with the problem of Malay morphology. Yet very few of these studies have touched upon the question of Malay expressives and the rules that relate to them. In fact, some authors have mistakenly identified certain expressives as 'verbal roots'.<sup>1</sup> The cause of these errors and oversights is unclear since expressives appear with considerable frequency in spoken Malay. The fact that they seldom appear in written Malay may explain why earlier writers, whose analyses were primarily textual, failed to concern themselves with these phenomena. Then, too, the standard languages of Malaysia and Indonesia, both variant forms of Malay, are essentially based on written Malay. Hence, modern studies of standard Malay likewise omit reference to this part of the language. Diffloth (1972) also observed:

The fact that ideophones are semantically unfamiliar (the majority of linguists alive today do not have ideophones in their respective mother tongues) and also difficult to observe for sociological and stylistic reasons can explain this neglect.

For whatever reason, analysis of the ideophones of Malay is sadly lacking. Some forms are cited in dictionaries, particularly those forms which are onomatopoeic. Studies of Malay dialects sometimes mention expressives, particularly with regard to the appearance of nasal vowels (Collins 1976; Wilding 1972; Brown 1927). Forty years ago there was considerable discussion regarding "the directional qualities and tones" of Malay sounds (Maxwell 1936). While there was spirited debate about the premise that all of the sounds of Malay words are based on onomatopoeia (Wilkinson 1936; Gonda 1940), no effort was made to isolate expressives from the rest of the vocabulary. Discussion about the role

of iconic language in Malay has been clouded because those forms which are exclusively iconic, expressives, have not been carefully analysed. While the expressive system of at least one other Austronesian language, Javanese, has been preliminarily defined and described (Uhlenbeck 1971), this is not the case for the expressives of Malay.

The assertion that "there is no clear-cut boundary between expressives and non-expressives" (Carr 1966) is not true. This paper hopes to provide an introductory description of expressives in Malay. The description is based on the Malay spoken in one area of West Malaysia, namely Kedah,<sup>2</sup> because, as has been mentioned, expressives rarely appear in the standard national languages of Malaysia and Indonesia. The paper is divided into five sections: a brief review of the theory about expressives and iconicity, analyses of Malay expressives and of their relations to what they signify, to other signs in the language system and to the speakers including some remarks regarding iconicity in Malay.

## 1. EXPRESSIVES AND ICONICITY

The category of words referred to here as expressive has been observed and noted for a long time. Brandstetter (1916:39-40) suggested that in Austronesian languages there were three kinds of "interjections": 1) those evoked by internal psychological processes; 2) those elicited by some external event; and 3) those which express a complete judgement. Those of the second group he further subdivided into words which are imitative of sudden events and those which are not imitative or not related to sudden events. It is this second group with all its subdivisions which constitutes the expressive system of a language. Brandstetter was right to distinguish these expressives on the basis of their relationships to real, external events and, yet, he was correct, too, to group them under the cover term "obscure formations of the emotional impulses". As Jakobson (1965) remarked, "An expressive signals the presence of certain sensations in the speaker."

Brandstetter's use of the term "interjection" for both exclamations and expressives is unfortunately misleading. There is an inherent connection between the expressive and the related event. This connection is one which exclamations do not have. Since Brandstetter's observation there has been some refinement of theory and terminology. Doke (1935) has proposed a semantic definition of the "ideophone"; others (Voeltz 1970, Baker 1938, 1939) suggest syntactic categories for them. There seems to be an accepted typology of the expressive (Johnson 1974). While there is some disagreement as to the scope of the relationship, most authors agree that expressives are in some sense imitative of the

event that evokes them. An expressive belongs to that kind of linguistic sign called icon. "An icon is a non-arbitrary intentional sign - that is, a designation which bears an intrinsic resemblance to the thing it designates" (Pierce 1931).

An expressive, then, conveys in a linguistic sign a sensate experience of sound, sight, touch or feeling in a way which resembles the external event which evokes the perception. There is a correspondence between the word and the perceived event or, rather, "between an acoustic image and a conceptual one, between phonetic and ideational values" (Friedrich 1976). As such, these words are distinct from the elements of language which are arbitrary or conventional correspondences of sound and meaning.<sup>3</sup>

Exactly how this relationship is effected remains obscure. Few experiments about the nature of iconic language have been undertaken (Wisseman 1954; Brown 1955; Markel and Hamp 1961). Current theories are based on data drawn from various languages and variant interpretations of these data. Often these data do not convincingly demonstrate the validity of any semantic or grammatical approach to expressives (Grossman 1975). At this point, this paper adopts the sound-symbolic approach which insists on the iconic relation of word and perceived event (Diffloth 1973). To what extent this position is justified with respect to the expressives of Malay will be considered later in this paper.

## 2. EXPRESSIVES AND THEIR SHAPES

As in many other languages, expressives in Malay can be recognised both by their peculiar phonetic qualities as well as by the existence of certain paradigmatic processes restricted to them. What these phonetic and organisational factors are will be discussed in this section.

There are at least two phonetic peculiarities of Malay expressives which mark them off from the rest of the vocabulary: nasal vowels and trilled [r]. It must be admitted that not all expressives involve either of these phenomena. Furthermore there are a few, very rare occurrences of both nasal vowels (in non-predicted environments) and trilled [r] in other parts of the lexicon.<sup>4</sup> While such phenomena are extremely rare in other parts of the lexicon, they are quite common in the expressive vocabulary.

In general Malay has no 'phonemic' nasal vowels; any nasalisation that takes place is due to the phonetic environment of the vowel. [Riŋɪt] '*dollar*'; [kəmɪʔ] '*dented*'; [m:ləh] '*to slaughter some animal*'

with proper Islamic ritual'; [akãɿ] 'root'. Nasalisation occurs immediately after nasal consonants /m, n, ŋ/ or after nasal plus liquid or glide. It also occurs preceding the pharyngeal fricative [ʕ].<sup>5</sup> In short, nasalisation is predictable - except for the nasal vowels of expressives.

[pɔ̃ʔ] 'of a stone tossed against a tree trunk'

[tĩt tĩt] 'of a chick when it is near its mother'

[siāp siāp] 'of moving very quickly like lightning'

[kuẽʔ kuẽʔ] 'of writhing movements (snakes in water, fish in mud)'

[biũ biũ] 'of swinging fists (that do not hit the mark)'

In these examples there are no conditioning environments to explain the appearance of nasal vowels. Here nasal vowels are an essential part of the expressives.

Malay spoken in Kedah has no trilled liquid; [R], an uvular fricative, occurs as well as its allophone in final position, [ʕ] (Asmah 1975 and 1977). In expressives, however, both [r] and [R] occur - particularly in words of onomatopoeic character.

[prep prep] 'of paper being crinkled and crushed'

[krcʔ kraʔ] 'of brittle objects knocking against each other like water buffaloes horn to horn, wooden floats of a net colliding with each other when a school of fish enters the net, stones in a coconut shell'

[greŋ greŋ greŋ] 'of a motor scooter warming up'

[braw braw] 'of heavy rain especially on tin roofs'

[praŋ] 'of a single piece of crockery falling on cement and shattering'

[Rop Rop Rop] 'of walking on dry moss in a fallow rice field'

[cRɛ:] 'of oil sizzling in a heated cauldron'

[sRop sRāp] 'of new batik brushing against itself when worn'

Something more will be said about the appearance of [r] and [R] in intervocalic position but at this point let it be noted that the appearance of [r] is highly unusual in the non-iconic vocabulary of Malay and yet rather frequent in iconic vocabulary.

Besides these two striking phonetic peculiarities of many expressives there are clear word-building processes for expressives; one might say 'morphology'. There are two main processes: affixation and reduplication including vowel and consonant alternation. These two processes often interact.

In the early part of this century Brandstetter mentioned some of the affixes associated with onomatopoeic words in Malay: [də-, kə-]. Both

Maxwell (1936) and Wilkinson (1936) have noted this affixation process and assigned functions to the affixes.

- (i) the sound: tok; bak; dam;
- (ii) making the sound: kětok; dēbak; lēbak; rēdam;
- (iii) repeating the sound: kělětok; kěrětok; gělēbak; kěrēdam.

(Wilkinson 1936:75)

Based on the data from Kedah it seems appropriate to sort out two, possibly three, kinds of affixes. There is an affix which indicates changed quality or intensity of the sound (or feeling) perceived - not "making the sound" as Wilkinson suggested. When it appears initially, this affix is [dē-]. It is prefixed to an expressive base.<sup>6</sup>

[nūt nūt] *'of repeated, 'sucking' pain as of an infected wound'*

[dēnūt] *'of a heart beat or pulse'*

[bum] *'of a ball being hit'*

[dēbum] *'of feet hitting the surface of the water while swimming'*

There is another affix which represents preliminary or auxiliary perceptions in addition to the main perception or, perhaps, repetition of that main perception. This affix appears in words that deal with sound perceptions only. When it appears immediately preceding the expressive bases, it is (usually) [kə-].<sup>7</sup>

[təŋ] *'of a metallic object falling'*

[kətəŋ] *'of cans being rattled against each other; of typewriter keys being hit too hard'*

[cāp] *'of lips smacking'*

[kēcāp] *'of repeatedly chomping on food'*

These two affixes, [kə-, dē-], can appear in the same word. The order invariably is /kə+dē+BASE/. This form represents repeated audial phenomena of varying sound quality often with accompanying sounds.<sup>8</sup>

[kədəkəŋ] *'of a very serious, probably fatal cough'* (Note: [kəŋ] *'of hitting someone's head'*)

When these affixes appear in combination, however, their phonetic appearance often varies: [kə- ~ gə-]; [dē- ~ rə- ~ lə- ~ tə-].<sup>9</sup> The form [gə-] always precedes the forms with affix [dē-] or, rarely, [rə-]. [-dē-] always precedes bases with initial voiced stops.

[gədəbaŋ] *'of a small drum; gunfire'*

[gədēbum] *'of a full bucket falling down in a well'*

[gədēbuʔ] *'of a coconut falling on (wet) sand; of any fruit falling'*

Before bases beginning with [p], [-lə-] or [-rə-] appear.

[kələpɔŋ] 'of an empty bucket placed on a stone'

[kələpuʔ] 'of thumping on wood'

[kərəpǎʔ] 'of a raga (ball made of plaited rattan) being kicked'

Before bases beginning with [k], any form may appear.

[kətəkɛŋ kətəkɛŋ] 'of metal struck to metal'

[kətəkam] 'of wood hitting wood (doors slamming)'

[kərəkāt] 'of horse-shoes tossed on cement'

[kədəkɔŋ] 'of a serious cough'

In addition to these affixes and combinations of them, there is possibly a third affix, the infix [Rvr].

[kiŋ] 'of a small stone tossed against a metal phone pole'

[kriŋ kriŋ] 'of coins shaken in a pocket; of a bicycle bell'

[baŋ] 'of a single piece of crockery falling'

[braŋ] 'of several dishes rattled together, shaken or set down hard; of thunder; of a net (with weights) tossed'

[sit] 'of deep drags taken from a cigarette'

[sRit] 'of a snake moving (in grass?)'

[cīt cīt] 'of a chick near its mother'

[cRīt cRīt] 'of repeated sobbing'

While informants seem to be able to isolate the function of [-r-] as an indication of a continuity of uneven sound,<sup>10</sup> it is unclear that this segment is an affix. Perhaps it is a meaning-bearing element in the composition of the expressive. Although with fairly complete accuracy we can predict that [r] will appear after non-continuant obstruents and [R] will appear after continuant obstruents, that is, there is seemingly allophonic predictability, it is not clear that the two have the same function. Brandstetter (1916:27) commented on a similar dilemma:

In Sundanese and Gayo there are interjections beginning with a mute and a liquid, e.g. Sund. drel, an interjection used of the rattling of musketry fire. At a pinch one might regard the -r- as the infix discussed in #86, in which case the root would only have three sounds, but the r seems to us so essential to the symbolic representation of the sound of rattling that we must decline on this occasion to take it for an infix ....

This comment as well as the data presented here suggests that within the expressive system the distinction between morphological paradigms and meaning-bearing elements is none too clear. It is hard to say that the 'morphological' affixes in the expressive system are not in themselves iconic. This point will be taken up later.

In addition to affixation, expressives undergo another change in form, reduplication. Some examples have already been cited. They include complete reduplication as well as partial reduplication with



changes in certain vowel and consonant segments. Reduplication indicates a repeated event or, in the case of repetition with different vowels and consonants, two events of different sensate quality. In some cases the expressive is repeated twice, perhaps to stress the continuity of the event.

Simple reduplication:

- [ciĕp ciĕp] 'of a chick far from its mother'
- [diʔ diʔ] 'of a steady, slowly falling rain; of a stiffening penis'
- [ŋɔŋ ŋɔŋ] 'of quick, straightforward strides'
- [sose sose] 'of spoken English'
- [s<sup>ə</sup>Raʋq s<sup>ə</sup>Raʋq] 'when brushing past leaves'
- [dəbum dəbum] 'of a dog swimming'
- [kətɛʔ kətɛʔ] 'of the waddle of a duck'
- [kələtiŋ kələtiŋ] 'of coins rattled in a bottle'
- [prɔʔ prɔʔ prɔʔ] 'of rattan cables being grasped while scaling a cave wall'

Reduplication with vowel change (and, in one case, [-r-] insertion):

- [klɛŋ klɔŋ] 'of Thai being spoken',<sup>11</sup>
- [tup tap tup tap] 'of a light rain on a tin roof'
- [bum bam bum bam] 'of flailing in the water, arms and legs raised above the water and coming down again and again',<sup>12</sup>
- [gədəbaŋ gədəbuŋ] 'of the rhythmic interplay of two drums of different size'
- [prup prəp] 'of wet pants cuffs'
- [kɔŋ kriŋ] 'of the small pestle and tube (gobek) used to soften betel preparations for toothless elders'

Reduplication with consonant change:

- [ŋũʔ ŋũt] 'of walking with a slow gait while nodding the head'
- [pro prit] 'of many different bird noises'

Reduplication with consonant and vowel change (and [-R-] insertion):

- [waŋ win] 'of oscillating objects (cradles, weaving drunks)'
- [cuʔ cRīt] 'of a broken-down bicycle with too little oil'
- [cuʔ cRāt] 'of walking in a flooded ricefield',<sup>13</sup>

Expressives display certain phonetic peculiarities and have a distinctive system of affixation and reduplication. In what way can these factors be related to the perceived events they represent? Do the sounds have a connection with the perceived experience?

## 3. EXPRESSIVES AND EVENTS

So far in describing the data I have followed van der Tuuk (1971) in attaching affixes to a base that is usually monosyllabic. But this approach obscures the role that each sound has in conveying a nuance of meaning. Many authors (including Brandstetter 1916, Wilkinson 1936, and Gonda 1949) have commented on the role of vowels in expressing different qualities: largeness/smallness, lightness/heaviness, etc. A few examples are given here.

- [tīt tīt] 'of a chick near its mother'
- [t3t t3t] 'of a telegraphic receiver'
- [piŋ] 'of a stone thrown at a telephone wire'
- [paŋ] 'of a slap'
- [pɔŋ] 'of a gendang (kind of drum) struck once'
- [puŋ] 'of gunfire'
- [prip] 'of a small bird (Ploceidae like sparrows and munias)'
- [prep] 'of paper being crinkled'
- [prap] 'of a school of fish swimming near the surface with tails rising in and out of the water'
- [prup] 'of a crisp food (fried fish, crackers, keropok) being eaten'

Similarly in reduplicated forms vowel difference is indicative of different events in combination or in sequence.

- [bum bam bum bam] 'of thrashing about in the water with arms and legs'
- [gədəbaŋ gədəbuŋ] 'of two drums, small and large, being played'
- [ŋɔŋ ŋɛŋ] 'of mosquitoes at ears; of a radio with static'
- [pūt pāt] 'of a repeated 'stabbing' pain, severe itchiness' (Compare to [pūt pūt] above)

Nasalsed vowels appear before stops when a louder, more resonant quality is signified. Nasalisation also seems to be associated with quick movements. Does its simultaneous superposition on other sounds convey speed?

- [wēt wēt] 'of repeated rapid, back and forth movements of the fingers'
- [siāp siāp] 'of the rapid movement of lightning'
- [cēt] 'of a mosquito bite or medical injection'

There can be no argument that vowel variation conveys a meaning: higher pitched sounds, more resonant sounds, more intense feeling. The vowels reflect a perceived event. Do consonants have similar correspondence to the events?

In final position consonants do seem to have some connection with the event. While nasals, fricatives and stops contrast in use, within each of these classes there is no clear distinction of use among the segments. Final nasals, predominantly /ŋ/ and secondarily /m/, are associated with resonant events. This is connected with the quality of the sound produced by nasal aperture. /m/ seems more associated with explosions and sounds in the water; whereas /ŋ/ appears with metallic, humming and drum sounds. The explanation for this seems unclear.<sup>14</sup>

Final stops /ʔ, t, p/ mark abrupt conclusion of sound or sensation in general, in some cases indicating contact. Again the articulatory process involved in the production of these sounds is one of contact and closure. To propose a connection of perceived event and sound seems acceptable.

Final [ɸ] occurs rarely but is always associated with light, hardly touching contact with friction as in the following:

[səRaʋɸ səRaʋɸ] *'of brushing against leaves'*

[buʋɸ buʋɸ] *'of a light breeze'*

Here too there seems a possible connection between the slight somewhat friction-filled contact in the event and the articulation.

Ø in final position is quite rare. The preceding vowel is invariably nasalised and often lengthened. Sounds which fade away are often represented this way; for example, [piũ piũ] *'of swinging fists'*.

Events with abrupt beginnings are described by words with stops in initial position. One can discern some meaning conveyed by choice of voiced or voiceless stop:

[plup] *'of a round object going into a space'*

[blup] *'of feet in and out of mud; of sexual intercourse'*

[pʃʔ] *'of a stone hitting a tree trunk'*

[bʃʔ] *'of a log hitting a wet mound'*

[praŋ] *'of a single dish fallen on cement and shattered'*

[braŋ] *'of several dishes rattled and shaken together but not broken'*

The voiced stop is associated with events involving greater resonance, a quality which distinguishes it from its voiceless counterpart.

Nasals appear initially often with onomatopoeic words describing humming sounds. But why are /n/ and /ŋ/ the initial sounds in words dealing with locomotion? Why is /p/ the initial sound in words dealing with kinds of pain?

[pʃŋ pʃŋ] *'of a straightforward gait'*

[ŋũʔ ŋũt] *'of a slow gait with a bobbing head'*

[pūt pūt] *'of repeated 'sucking' pain'*

It is not clear that this choice of /p/ and /ŋ/ is iconic.<sup>15</sup>

Initial frictional continuants /s, c/ mark the beginning of a brushing movement where some contact is made with some friction.

[cu' cRāt cu' cRāt] *'of the contact between feet and mud while walking through flooded ricefields'*

[kēcāp kēcāp] *'of repeatedly chomping on food'*

[cRēt] *'of something being crushed'*

[sRop sRap] *'of new batik brushing against itself when worn'*

The same applies to initial /R/, that is, some contact with friction.

[Rot Rīt] *'of a bicycle without oil'*

[Rop Rop Rop] *'of walking on sand or the dry moss of a fallow ricefield'*

In these cases then there seems to be some perceived connection between the abrasive contact and the sounds made with friction.

Initial /w/, a sound made with both lips, is found in words describing back and forth or oscillating movements.

[weō weō] *'of things or persons moving back and forth rapidly'*

[waŋ win] *'of oscillating movements (like cradles, drunks, etc.)'*

[wēt wēt] *'repeated, rapid back and forth movement of the fingers'*

In earlier parts of the paper attention has been drawn to the role of medial sounds, that is liquids. /r/ and /l/ both seem to have iconic functions in the words they appear in (see p.384). Furthermore, there are indications that the choice of [d, l, r, t] in affixation may also be related to iconicity (see note 9).

In general the sounds which appear in expressives seem to be related to the experienced events they describe. Similarly the sequence of sounds within an expressive is often (?) iconic.

[cu' cRāt] *'of walking in a flooded ricefield'*

[c], initial abrasive contact with the water;

[u], resonance as foot sinks into the mud;

[ʔ], abrupt halt as foot reaches firmer bottom;

[c], friction as foot is withdrawn from clinging mud;

[R], continued friction with mud;

[ā], resonant quality as foot is released from mud (with suction?);

[t], abrupt conclusion as foot is free of mud and water.

[prīt] *'of cloth being torn'*

[p], initial tear of cloth;

[r], continued ripping;

[i], with a high pitched sound (this sound is different when the cloth is wet);

[t], abrupt conclusion as cloth is torn through.

(Note this expressive also refers to spitting in small globules, urinating in dribbles and the sound of a referee's whistle)

[braw braw] *'of a heavy rain on a tin roof'*

[b], initial contact of rain on a tin roof;

[r], persistent roaring noise as rain increases in intensity;

[aw], resonant quality.

In non-iconic language, sounds, word-shapes and word-order are distinguishable. In iconic language the word is meant to portray a perceived event (or, better, one's reaction or interpretation of an event). In such a word, each sound bears a meaning. The choice of sounds is not arbitrary; each sound as a relationship to the perceived event. As such it is difficult to distinguish possible 'morphemes' from 'phonemes'; note the dilemma regarding infixed /r/.

The arrangement of sounds within an iconic word is related to a perceived series of events. Neither the individual sounds nor the order in which they are arranged is arbitrary. On the other hand, neither are they exact duplicates of the event. There is scope for variety among the speakers within some range of conformity. More will be said on this subject in section 5 of this paper.

While it is convenient to use a monosyllable stem or base as a starting point for the description of expressives, in fact, each segment in that base conveys a meaning about the perceived event. In most cases that communication is iconic. In some cases the iconic connection is not transparent. It could be that in those cases the connection is conventional.<sup>16</sup> It could be equally true that there is an iconic relationship. The production of sounds is a complex process involving many physical movements and many internal sense experiences. Some features of those movements and sensations may be selected as the acoustic parallel of the perceived event. While there is no evidence to support either possibility, the fact that most of the phonetic segments of expressives in Malay are immediately, patently iconic lends some weight to the possibility that all of the segments are iconic, perhaps in ways not so immediately obvious to the non-native speaker. In short, there is considerable evidence to support the notion that sound-symbolism is central in expressive language.

## 4. EXPRESSIVES WITHIN THE SYSTEM

In the course of this paper, the terms 'iconic language' and 'non-iconic' or 'logical' language have been used almost as if they represented two different languages. Of course this is not the case. Human language seems to be organised along a spectrum; there is a range spreading from the expressive and emotive to the rational and notional (Brandstetter 1916; Carr 1966). Although iconic and logical elements are different from each other, they seem to complement each other. It seems safe to say that no language is without either component and there are shadings and gradations between them. European literary languages, for various socio-cultural reasons, display minimal use of expressives; whereas Bantu, Korean and Aslian languages seem particularly rich in expressives. The point to be stressed here is that expressives are part of the total language system. As such they share resemblances with the non-iconic element and they interact with that element to form the basis for communication.

Certain phonetic peculiarities have been noted in Malay expressives. It was admitted, though, that while the frequency of occurrence of these peculiarities was unusually high within the expressive component, such peculiarities do occur in the rest of the language, albeit extremely rarely. The range of sounds found in expressives is contained within the total 'phonemic' inventory of the language. Furthermore the constraint on consonant clusters other than occlusive plus liquid that is apparent in expressives is one which operates in the non-expressive items of Malay.

Similarly it seems likely that certain sound changes which occur in Kedah Malay in general also occur within the expressive system. Final /s/ reconstructed for many Austronesian words and appearing in most dialects of Malay appears in Kedah Malay with the phonetic form [ɤ]. (Other authors, for example Asmah 1977, treat this sound differently.)

One could postulate a rule, perhaps diachronic, to explain this divergence:

$$(1) \begin{bmatrix} +\text{continuant} \\ +\text{strident} \\ +\text{anterior} \end{bmatrix} \rightarrow [-\text{anterior}] / \_\_\_\#$$

For example: \*halus → [haluɤ] 'fine, delicate'

\*tirus → [tiRuɤ] 'tapered (especially of tails)'

Expressives like [buɤ buɤ] ('of a gentle breeze') probably have undergone this same rule. (In other dialects not subject to this sound change the cognate form is [bus bus].)

In Kedah Malay as well as in other dialects, namely Patani and Kelantan Malay (Wilding 1972; Asmah 1975), final stops converge to [ʔ].

(2) [-continuant] + Ø [+glottal] \_\_\_\_ #

For example: /kəjaʔp/ → [kəjaʔ] 'firm, fixed'

\*kələntiʔt → [kələntiʔ] 'clitoris'

Expressives like [cẽʔ ~ cẽt] ('of an injection or a mosquito bite') as well as [ɲuʔ ɲut] ('of walking with a slowly bobbing head') probably reflect this sound change. Just as in the rest of the language the change from glottalised occlusive to glottal occlusive is not universal and represents a direction the language is heading.<sup>17</sup>

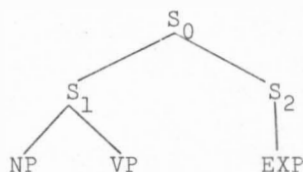
While the affixes used with expressives are unique within the language system,<sup>18</sup> they too follow the restriction imposed on all affixation in Kedah Malay. There are no suffixes. All affixation is with prefixes. In reduplication which involves change in consonant form, the change will take place in the first occurrence of the reduplicated item.

Expressives, then, are subject to the limitations and constraints of the total language system. They must also interact within that system. They must occur within or with the syntax of the language. Most often expressives appear immediately preceding or following the sentence.

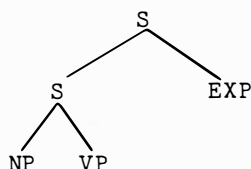
- (1) [hujan tuRun braw braw]  
rain, descend, 'of a heavy rain on a tin roof'  
'The rain fell with a tremendous uproar.'
- (2) [pələmpɔŋ bəlaɣa kruʔ krãʔ]  
float, MID+behave, 'of brittle things knocking together'  
'The floats (on the net) rattled against each other.'
- (3) [bisui aku sakit put put]  
boil, I, sick, 'of a steady 'sucking' pain'  
'My boil is throbbing with pain.'
- (4) [səRuʔq səRuʔq aku tariʔ kayn di ataɣ pasiaʔ]  
'of something dragged', I, pull, cloth, at, above, sand  
'I was dragging some cloth across the sand.'
- (5) [kuaʔ kədəɣay kutip anaʔ]  
'of hasty collection of objects', collect, offspring  
'(We) collected the young (birds) one after the other.'

In these sentences the expressive is in a sort of appositional or, better, equational relation with the logical statement. The two statements, iconic and notional, seem to say the same thing. The one explains the other. In (2) the first part tells exactly what happened: 'floats

*were moving*'; the second part expresses the rattling sound as well as the turbulent, erratic quality of that moving. A diagram of their syntactic relationship might be:



This implies that the expressive itself is a sentence. Maybe this is so but just as the labels 'phoneme' and 'morpheme' have dubious value in describing expressives the label 'sentence' is equally misleading. While it is true that an entire concept is being expressed, it is not exactly clear what is being predicated of what. Only circumstances can make that clear. This will become more apparent when speaker attitudes are discussed. For the time being it seems best to present as a tree diagram the following:



We are reluctant to specify a label for the node dominating the expressive.

Sometimes the expressives occur after the word /bupi/ ('sound') or /kəna/ ('experience (something unpleasant)').

- (6) [buah macaŋ luRuh bupi gədəbu?]  
 fruit, mango, drop, sound, 'of a fruit falling through branches to the ground'

'A mango dropped through the leaves to the ground.'

- (7) [kəmaRin lan kəna piʔäp]  
 yesterday, Lan, experience, 'of a rotan cane passing through the air'  
 'Yesterday Lan got it with the cane.'

Note that in (6) it is possible to delete /bupi/ and the sentence remains acceptable. /bupi/ then functions as a sort of connector<sup>19</sup> between iconic and logical statements. It is with this connector that some expressives make their way into written Malay.<sup>20</sup> In (7) the situation is different. In this sentence the format is perhaps closer to that suggested for (2). The sentence [kəmaRin lan kəna] ('yesterday, Lan, experience') can usually stand alone with the meaning 'Yesterday Lan got it.' [piʔäp] serves to express the quality of that unfortunate



experience with a word quite vivid, emotional and specific for most Malay schoolboys.

There are some instances where an expressive occurs in the middle of a sentence.

- (8) [bila dia pa:ŋ kəna baRu lah təkəpah kəpah təkətaŋ kətaŋ]  
 when, III, 'of a slap', experience, new, EMPH, INAD+shake-shake,  
 INAD+shiver-shiver

*'When he gets it good, then he'll shake and shiver with fear.'*

Certainly in meaning this construction is similar to (7) only the expressive has been placed before the word /kəna/ ('experience'). This may represent a transformation whereby the expressive is optionally moved closer to a relevant item for stylistic purposes. It is possible to postulate a further transformation for the following sentence. Speaking of a minor scrape a car received,

- (9) [dia cẽʔ]  
 III, 'of slight abrasive contact'  
*'It just nicked (it).'*

One might suppose a series of transformations:

- \*/dia kəna cẽʔ/  
 \*/dia cẽʔ kəna/ I. 'expressive placement'  
 /dia cẽʔ/ II. '/kəna/ deletion'

On the other hand, it may not be necessary to propose a hypothetical 'standard' deep structure for sentences with expressives. Since expressives seem to be outside the scope of a sentence they can be fitted in where they seem to the speaker to be appropriate. This kind of analysis is pleasing but it does not account for the tendency of expressives to become adjectival.

In addition to sentences like (9) where, without the analysis proposed here, it would be difficult to classify [cẽʔ] as other than verbal, sentences like (10) also occur.

- (10) [apa bunī prit prit tu]  
 what, sound, 'of ripping cloth', that  
*'What's that ripping noise?'*

Here the appearance of the deictic [tu] provides somewhat convincing evidence that the expressive has become a verbal or functions as a verbal. The fact that the affixes which indicate verbal relationships are often not used in spoken Malay facilitates the movement of expressives into the logical system of language.<sup>21</sup>

- (11) [aku pi pantay anin may buʔq buʔq]  
 I, go, seashore, wind, come, 'of a light breeze'  
*'I went to the beach; the wind blew gently.'*

In this sentence it is difficult to determine if the expressive stands in an equational relation to the sentence /aŋin may/ or if it is a kind of verbal. This case with which expressives fit into the notional patterns of the language is all the more apparent when expressives appear with verbal affixes.

- (12) [dia sakay jit sampay məŋəRʂh]  
 III, *strike violently, Aziz, reach*, INTR+'of a low rumbling like snoring, groaning and distant thunder'  
 'He beat Aziz until he groaned.'
- (13) [anaʔ ayam tu bæCRĕt]  
 offspring, chicken, that, MID+'of something animate being crushed'  
 'That chick was crushed.'
- (14) [ulaŋ tu təkuĕʔ kuĕʔ jalan di aiăŋ]  
 snake, that, INAD-'of writhing movements', move, at, water  
 'That snake slithered away in the water.'

In these sentences the expressives appear marked by the peculiar characteristic of vowel nasalisation but they are affixed by the verbal markers: /məN-/ for intransitive verbs; /bə(R)-/ for middle voice; /tə(R)-/ for inadvertant actions. The expressive has been drawn into the notional system of language. It may happen that such an expressive loses its vivid force for the speakers. In (15) this may be the case.

- (15) [buRɔŋ tu kukʂŋ krʂʔ krʂʔ]  
 bird, that, coo, 'sound of raspy bird noises'  
 'The bird cooed, cooed and trilled.'

Apparently kukur ([kukʂŋ]) was not sufficiently vivid for the speaker so he felt the need to add a fresh, patent expressive.<sup>22</sup> On the other hand, this may be a string of expressives, the one more specific than the other.

In general it may be said that expressives function in apposition to the logical description of the event which they describe. In some cases the expressive is drawn into the elastic, 'omnivorous' Malay verbal system. While retaining its iconic force, it assumes verbal affixes and a verbal function within certain sentences. Since language includes a spectrum of emotive and notional components, it should not be surprising that there are phenomena that seem to partake of both components. It seems likely that the Malay expressive for the sound of spoken Thai [kleŋ kleŋ] developed from [kleŋ] the Thai word (southern dialect) for 'language'. Few, if any, monolingual speakers of Malay would be aware of that connection. Hence, it seems there is a two-way movement between the two poles of language. This movement and the failure to perceive its dynamic, unstable character is at the root of the unacceptable explanations of Malay word-building found in C.N. Maxwell's writings.

## 5. EXPRESSIVES AND EXPRESSERS

In this section, a brief summary of the attitudes of the speakers of Malay towards expressives is noted. Most speakers have definite opinions about the relationship of expressives and the perceived events they portray. Some speakers have commented on the position of expressives in the total language system. Perhaps following the clues of these informant attitudes we can evaluate the findings of the earlier sections.

There was a consensus that the words had a clear connection with the perceived event. In the case of onomatopoeia, the words were said to be based on the sound of the event. In fact, one should use the expressive only at the moment the event occurs. One informant was certain that it was incorrect to use the word out of context or to use it in a circumstance that was inappropriate to the expressive. All agreed there was some difference of interpretation both of the perceived event and of the way to express it. Animal noises were thought to be clear-cut; they speak for themselves and so only certain orders of sounds can represent them. The facts, however, seem to bear out the comment of one young informant who said that the sounds of the expressive are relevant at the moment of reaction and that each person's reactions are different.

This became increasingly clear in the course of the study. When asked to identify a sound one speaker would offer one expressive; in another instance another speaker might offer another expressive.

[diʔ diʔ]  
[tup tap tup tap] } 'of a light rain'

If a speaker spontaneously inserted an expressive in his speech and was asked to identify exactly what the expressive referred to, there often was denial that the expressive was ever used. Otherwise there was considerable reflection as to what was being captured by the word. When that word was presented to another speaker and he was asked what it referred to - a common enough procedure for other vocabulary items - the response would sometimes be that the other person invented the word. At other times the speaker would give a meaning almost exactly the same as the first speaker. For example:

[cu' cRit] Speaker 1: 'of an old bicycle'  
Speaker 2: 'of a bicycle that needs oil'

On other occasions an expressive might elicit an entirely different meaning. For example:

[di' di'] Speaker 1: *'of a light rain fall'*

Speaker 2: *'of a slowly stiffening penis'*

When further cross-checked with other speakers, the response would often be: "Of course, that just the word for it! What a fine way to express it!" The creative speaker was admired for his ability to catch some nuance of sensation with a few appropriate sounds. Some speakers were noted as being particularly prone to use expressives or particularly good at using them. On other occasions people would discount another speaker's use of an expressive for a certain circumstance as being inconsistent with the actual sounds or perceptions.

The general impression is that there is a broad range of interpretation and inventiveness. Some expressives are rather commonplace and heard often; [pru'] *'of a rattling noise'* is used frequently and is coming to mean *'place something down coarsely'*. Other expressives seem to be part of an expressive vocabulary that is rather stable. But this may not be the case for all expressives. Some seem to be passing out of the range of acceptability. A recorded narrative of an 80 year old man was punctuated by many expressives. One of them caused confusion and later laughter when played for younger informants.<sup>23</sup> Some expressives are new, spontaneous inventions created on the iconic principles outlined in section 3. There are criteria for judging the appropriateness of a given expressive to a given circumstance but everyone admits that it is still dependent on how the speaker feels - that is, what aspect of the event he focuses on.

Expressives are recognised as a class. Speakers have variously referred to them as *bahasa runcit*, *bahasa serpihan* and *bahasa saduran*. These reflect three different qualities of the expressive system. The first is that expressives exist in great quantity; there is an unlimited supply of many different kinds of these words. There will always be one suited to the particular needs of a circumstance.<sup>24</sup> The second phrase refers to another quality of iconic language. Expressives are fragments. They are not complete in themselves. In Malay expressives obligatorily appear with logical sentences which specify the locus of the emotions conveyed. Expressives do not constitute a complete language system in themselves. This is further evidenced by the third quality captured in these phrases. Expressives are the gilding of the total language system. While they do not stand by themselves, they constitute the colourful, rich coating on the surface of logical language. They contribute a vivid, sparkling quality to spoken sentences and, in that sense, sentences too can not stand apart. Iconic and notional language are fused together and inseparably important to each other.

## N O T E S

I have to say thank you to Gérard Diffloth who first drew my attention to expressives over three years ago. He also provided an informal, informative atmosphere in his lectures on Semai expressives. The research upon which this paper is based took place on Langkawi, Malaysia. The people there showed their good nature and sense of humour as I sorted out their language. Osman Mahmud was especially patient. A note on symbols: final stops in Kedah Malay are glottalised; in this paper that glottalisation is not ordinarily indicated. Although standard Malay orthography inserts /ə/ e between occlusive and liquid, that insertion seems predictable and is not noted in the paper except in certain verbal affixes.

1. In Abdullah Hassan's work (1974:237-8) one can find: debap, decit, degam. These 'verbal roots' come from iconic bases: /bap, cit, gam/ respectively.

2. More specifically the data are drawn from fieldwork done in Padang Matsirat, Langkawi. There are some limitations to the use of only the material I collected there in a two-three month period. The amount of expressives noted is rather small and mostly representative of audial sensation. These may reflect the language of a small subgroup of Malay speakers so the use of Malay in the title and elsewhere is questionable. However, to my knowledge there is no reliable source for the study of expressives in Malay. Dictionaries do not ordinarily mark them as a class and sometimes offer misleading information. I tried to verify some of these expressives among other speakers of Kedah Malay and Malay dialects in other parts of the peninsula. I was amazed at the number of cases of correspondence of meaning for words that have never found their way into a dictionary.

3. Friedrich (1976) distinguishes between conventional and arbitrary signs. Conventional signs cannot be said to be arbitrary.

4. In Collins 1976 I cite a few of these instances. /siã/ 'small, stackable containers for transporting food'; /trit/ 'the thread of a screw'. These borrowings as well as cross dialectal borrowings, e.g. /goren/ ('fry'), sometimes reflect these phonetic peculiarities.

5. See Abdullah Hassan 1967 for a complete discussion of predictable rasalisation in Kedah Malay.

6. In his study of Toba Batak van der Tuuk (1971) suggested that in certain onomatopoeia words the form originated from a monosyllabic stem preceded by h or g with the infixation of um and ar. It is interesting to note that many forms similar to those of Batak appear in Malay dictionaries, for example, gemercing, gemerlap, etc. Forms like these do not occur in Kedah Malay.

7. There is one recorded instance of /gədum gədum/. Choice of /g/ over /k/ may be dictated by the following voiced stop of dum.

8. Informants have stressed the presence of auxiliary noises. Fruit falls with a thud (/bu?/) but it is preceded by snapping leaves and branches so /gədəbu?/.

9. Because of the limited data it seems impossible to determine if these variations can be predicted on some clear phonetic grounds in all cases. Siti Hawa (1970:243) notes that /k/ and /g/ often are interchangeable in Kedah Malay; some of the forms she cites are expressives (gerisik: kerisik). Note too that /də, tə, lə, rə/ share the same articulatory position. Manner of articulation differs. It is possible that choice of the specific consonant of this affix has an iconic (and therefore somewhat idiosyncratic) underpinning. /r/ is for rattling noises like horse-shoes on pavement; /t/ is for abrupt noises like hard objects hitting each other; /d/ for resonant staccato noises like a serious cough. Two forces intersect here: the one phonetic assimilation, e.g. voiced stop before voiced stop; the other iconicity, e.g. trills for rattling noises.

10. He actually said bunyi yang bergulung-gulung, sounds which continuously roll.

11. /kleŋ/ in the dialect of Thai spoken in certain villages on the north coast of Langkawi means 'language'.

12. In this word the initial base /bum/ is repeated with vowel change to indicate changed audial quality. Then that reduplication is reduplicated to show repeated action. Informants insisted that the variation reflected the difference in sound created by raising arms and legs in the water. This seems reasonable compared to other forms cited in the paper, namely /bum, dəbum, dəbum dəbum/.

13. Here again the informant was insistant that /cuʔ/ represented the step down and /cRāt/ the movement of the foot upward.

14. Interestingly enough the tendency of the dialect is for final N to be realised as /n/ rather than /m, ŋ/. Is this another peculiarity of the expressive system?

15. Grossman (1975:35) suggested that the resonance created in the nasal cavity by /ŋ/ is "an oral counterpart to the physical quality of movement or rippling of shape". Perhaps it is this factor which is involved in selecting nasals for the description of locomotion. It might be possible to consider that the slight movement involved in producing /p/ is associated with the onset of pain.

16. See the results of Markel and Hamp's (1961) experiments.

17. These two rules might be more elegantly represented in Ladefoged's (1971) feature system. (A binary notation is used here rather than a multi-valued notation; this divergence is not too significant.)

(2)  $\left[ \begin{array}{l} +\text{stop} \\ -\text{back} \\ +\text{glottalic stricture} \end{array} \right] \rightarrow [+back] / \_\_\#$

(1)  $\left[ \begin{array}{l} +\text{sibilance} \\ +\text{tension} \end{array} \right] \rightarrow [-\text{tension}] / \_\_\#$

In particular the change from glottalised stop to glottal stop is better represented this way because then this is a case of coarticulation not a case of a sequence of two articulations as might be inferred from the notation on p.391.

18. /kə-/ does in fact occur as both a verbal affix and a nominal affix. However in both cases it is ordinarily accompanied by the suffix /-an/. Of course, these remarks are limited to standard Malay. While

I have not recorded any occurrences of verbal /kə-/ in Kedah Malay, there are a few cases of the nominal /kə-/. In these cases it occurs without the suffix /-an/, for example, /kələŋka?/ 'equipment'.

19. Grossman (1975) might say 'frame' as in the English sentence: 'The wind went whoosh.' Here 'went' is considered a frame for the onomatopoeic expressive 'whoosh'.

20. Zainu'l Ahmad (1927) cites a number of cases especially p.330.

21. But even here note that it is possible to move the expressive out of the apparent NP. So, [apa prit prit buni itu] and [apa buni itu prit prit] are acceptable. Verbals cannot be moved in a similar fashion.

[apa buni bəsaʔ tu]  
what, sound, big, that  
'What was that loud noise?'

This sentence cannot be similarly shuffled: \*[apa buni itu bəsaʔ] and \*[bəsaʔ apa buni tu] are clearly not acceptable.

Furthermore evidence exists that the expressives are not absorbed into the verbal system. They cannot be negated with the verbal negative: tidak, dak. If negation is possible at all, the non-verbal negative is used: bukan. Even in this case, though, negation is only possible as a comment on someone else's choice of expressives. For example:

- A. [buah macaŋ luRuh buni gədəbuʔ] (cf. #6)  
B. [bukan gədəbuʔ; dia buni laŋn, puʔ]  
NEG, 'of falling fruit', III, sound, other, 'of falling fruit'  
'Not gedebuk! It sounded different: puk.'

What is being negated is another's opinion, the preceding statement of the discourse. The scope of the negation is not the expressive itself but the entire statement. It is impossible to open a discourse with a sentence like:

(6a) [buah macaŋ luRuh bukan (buni) gədəbuʔ] (cf. #6)

Expressives themselves are not subject to negation either as verbals (with [tidaʔ]) or non-verbals. What is their status then in the language?

22. If it is true that kukur lost its vividness for that speaker it might be related to the sound change that final /r/ undergoes in Kedah Malay, \*r → ʕ / \_\_\_\_#. The effect of the trill or fricative is lost when the sound pharyngealises.



23. The expressive /kuaʔ kədəgay/ was recorded in a narrative about collecting edible birds' nests in neighbouring islands (see sentence (5)). This industry was common some 40-50 years ago. When this narrative was played for highschool students in the village, some could not identify it as an expressive at all and most were amused. It is likely that this expressive belongs to an older 'genre' of expressives. Or perhaps it was an idiosyncratic creation which was not well-received. In any case the point remains the iconic base that led the old man to use/create this expressive was not the same basis that would be acceptable to speakers 60-70 years his junior.

24. Since *runcit* could have an overtone of petty and unimportant, I specifically asked for clarification. He said he did not consider those words trivial, only exceedingly numerous.

BIBLIOGRAPHY

ABDULLAH HASSAN

- 1966      *Perbandingan tatabunyi antara dialek Kedah dan dialek Perak*. Latihan ilmiah (S.M.) Universiti Malaya.
- 1974      *The Morphology of Malay*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

ASMAH HAJI OMAR

- 1975      Language in Minority/Majority Group Relations: The Case of the Diversity of the Malay Dialects. Paper read at the Workshop on Linguistic Problems in Minority/Majority Relations, Bangkok.
- 1977      *The Phonological Diversity of the Malay Dialects*. Kuala Lumpur: Dewan Bahasa dan Pustaka.

EAKER, J.A.

- 1938      'Notes on the Meanings of some Malay Words II'. *JMBRAS* 16/2:47-50.
- 1939      'Notes on the Meanings of some Malay Words, Part III. (Kedah Words)'. *JMBRAS* 17/1:107-20.

BLAKE, F.R.

- 1977      'Reduplication in Tagalog'. *American Journal of Philology* 38:425-31.

BOLINGER, D.L.

- 1950      'Rime, Assonance, and Morpheme Analysis'. *Word* 6:117-36.

## BRANDSTETTER, R.

- 1916      *An Introduction to Indonesian Linguistics*. (Translated by O. Blagden). London: Royal Asiatic Society.
- 1917      *Die Reduplication in den indonesischen, indiaischen und indogermanischen Sprachen*. Luzern.

## BROWN, C.C.

- 1927      *Kelantan Malay*. Singapore: Government Printing Office.

## BROWN, R.W.

- 1955      Review of *Untersuchungen zur Onomatopoeie ...* by Heinz Wisseman. *Language* 31:84-91.

## CARR, Denzel

- 1966      'Homorganicity in Malay/Indonesian in Expressives and Quasi Expressives'. *Language* 42:370-7.

## COLLINS, J.T.

- 1976      'Vokal sengau di dalam bahasa Melayu Kedah (Langkaw1)'. *Dewan Bahasa* 20:19-31.

## DIFFLOTH, G.

- 1972      'Notes on Expressive Meaning'. *CLS* 8:440-7.
- 1973      Expressives in Semai. Paper read at the First International Conference on Austroasiatic Linguistics, Hawaii.

## DIMOCK, E.C.

- 1957      'Symbolic Forms in Bengali'. *Bulletin of the Deccan College Research Institute* 18:22-9.

## DOKE, C.M.

- 1935      *Bantu Linguistic Terminology*. London.

## EMENEAU, M.B.

- 1969      'Onomatopoeitics in the Indian Linguistic Area'. *Language* 45:274-99.

## FRIEDRICH, P.

- 1976      'The Symbol and its Relative Non-arbitrariness'.

## GONDA, J.

- 1940 'Some Remarks on Onomatopoeia, Sound-symbolism and Word Formation à propos of the Theories of C.N. Maxwell'. *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 80: 133-210.
- 1949 'The Functions of Word Duplication in Indonesian Languages'. *Lingua* 2/2:170-97.

## GOUDGE, T.

- 1950 *The Thought of C.S. Pierce*. Toronto.

## GROSSMAN, R.

- 1975 *Bahar Expressives*. M.A. essay, University of Chicago.

## HAAS, M.R.

- 1942 'Types of Reduplication in Thai (with some Comparisons and Contrasts taken from English)'. *Studies in Linguistics* 1/4:1-6.

## HAMILTON, A.W.

- 1922 'Penang Malay'. *JSBRAS* 85:67-96.
- 1925 'Corrigenda'. *JMBRAS* 3/3:56.

## IDA AHMAD

- 1969 'Fonoloji bahasa Kedah'. *Dewan Bahasa* 13:531-7.

## JAKOBSON, R.

- 1965 'Quest for the Essence of Language'. *Selected Writings* II, 345-59. The Hague: Mouton

## JOHNSON, M.

- 1973 *Towards a Definition of the Ideophone in Bantu*. MS.
- 1974 'Phonetic Play in Zulu'. *CLS* 10:

## LADEFOGED, P.

- 1971 *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.

LOPEZ, C.

- 1950 'Reduplication in Tagalog'. *BijdrTUV* 106:151-311.

MARKEL, N.N. and E.P. HAMP

- 1961 'Connotative Meanings of Certain Phoneme Sequences'.  
*Studies in Linguistics* 15/3-4:47-61.

MAXWELL, C.N.

- 1932 *The Malay Language and How to Learn it*. Kuala Lumpur.
- 1933 *An Introduction to the Elements of the Malay Language*.  
Kuala Lumpur.
- 1936 'Light in the Malay Language'. *JMBRAS* 14/3:89-154.
- 1938 'Language Affinities'. *JMBRAS* 16/1:1-99.

NOR EIN MOHD NOOR

- 1977 'Onomatopia dalam bahasa Malaysia'. *Dewan Bahasa* 4:253-63.

PIERCE, C.S.

- 1931 *Collected Papers*, vol.2: *Elements of Logic*. Cambridge,  
Mass.: Harvard University Press.

SAPIR, E.

- 1921 *Language: An Introduction to the Study of Speech*. New  
York: Harcourt, Brall.

SITI HAWA SALEH

- 1970 *Hikayat Merong Mahawangsa*. Kuala Lumpur: University of  
Malaya.

THOMPSON, M.H.

- 1953 *The Pragmatic Philosophy of C.S. Pierce*. Chicago: Univer-  
sity of Chicago Press.

THUN, N.

- 1963 *Reduplicative Words in English*. Dissertation, Uppsala.

TUUK, H.N. van der

- 1971 *A Grammar of Toba Batak*. The Hague: Nijhoff. (Original:  
*Tobasche Spraak[k]unst*, 2 vols, 1864, 1867.

## UHLENBECK, E.M.

- 1971 'Peripheral Verb Categories with Emotive-expressive or Onomatopoeic Value in Modern Javanese'. *Travaux Linguistiques de Prague* 4:145-56.

## VOELTZ, E.F.K.

- 1970 'Toward the Syntax of Ideophones in Zulu'. In: C-W. Kim and H.F.W. Stahlke, eds *Studies in African Languages and Linguistics*. Champaign, Canada: Linguistic Research Inc.

## WESCOTT, R.W.

- 1971 'Linguistic Iconism'. *Language* 47:416-28.

## WILDING, A.

- 1972 *Introducing Pattani Malay, Part B: Outline of Pattani Malay Sounds*. Yala.

## WILKINSON, R.J.

- 1936 'Onomatopoeia in Malay'. *JMBRAS* 14/3:72-88.

## WISSEMAN, H.

- 1954 *Untersuchungen zur Onomatopoeie 1: Die sprachpsychologischen Versuche*. Heidelberg: C. Winter.

## ZAINU'L-ABIDIN BIN AHMAD

- 1927 'Pair-words in Malay'. *JMBRAS* 5/2:324-38.

THE SAMOAN CONNECTION  
OR  
VERB AND NOUN PHRASE RELATIONS IN A POLYNESIAN LANGUAGE

MICHAEL R. THOMAS

INTRODUCTION

The question of the role of the nominal particles in Samoan has been discussed frequently since Churchward's innovative study in 1926. In that work Churchward differed with earlier missionary and trader interpretations of the verb phrase and associated subject noun phrase as being marked active or passive (instead considering them transitive and intransitive), but most subsequent studies have continued to follow the active/passive dichotomy. Regardless of which interpretation is used, however, the nominal particles have required a certain amount of imaginative definition. It is the purpose of this study to examine previous descriptions of these particles, and to suggest that a focus-type analysis is not only possible for a Polynesian language, but may result in a model of greater explanatory power.

- (1) na    mana'o le tinā    'i    lana tama.  
tense want    the mother part. her child
- (2) na    nana'o-m-ia le tama e    lona tinā.  
tense want    sfx. the child part. her mother

Both (1) and (2) may be translated '*The mother wanted her child*', or '*The child was wanted by her mother*' (Milner 1962:152). The traditional interpretation is to assign the active translation to (1) and the passive to (2). This results in the identification of the particle 'i as a marker for the direct object noun phrase, and the e as the preposition 'by'. In addition, the verbal suffix in (2) was considered to mark the passive form of the verb. With the examination of more sentences, however, the above description has had to be modified *ad*

*infinitum* to account for the variability in the particles' distribution and interpretation.

George Milner has questioned the descriptive adequacy of the above model in articles pointing out their three major flaws (Milner 1962, 1974). For example, sentences with passive suffixes such as (2) may be made 'passive in form, but active in meaning' by preposing the pronoun (as in (2a)). The pronoun loses its nominal particle in the process and superficially resembles an active sentence in which a pronoun has also been preposed (1a).

(1a) na ia mana'o 'i lona tama.  
t pn want part. her child

(2a) na ia mana'o-m-ia le tama.  
t pn want sfx. the child

Additionally, in some sentences the passive actor is not marked by the particle *e*, but is instead introduced like the object of an active sentence.

(3) 'ua ia lavea i lona uso.  
t pn hurt part. her sibling  
'She is hurt by her sibling.'

Thirdly, there are sentences which are described as passive although their verbs are not marked by a passive suffix.

(4) na fasi le tama e lona tinā.<sup>1</sup>  
t beat the child part. her mother  
'The child was beaten by her mother.'

In an attempt to eliminate some of the inconsistencies in previous descriptions of Samoan syntax, Milner suggested that it was aspect and not voice which was marked by the verbal suffix. Further, he divided the sentences into two basic types: Objective and Ergative.

Objective sentences are described as those in which: (1) the unmarked NP is identified as the actor and is the subject of a focused verb, (2) the NP introduced by the particle *i* is identified as the goal or object NP, and (3) the unsuffixed verb is in the imperfective aspect which stresses the action or process itself. The ergative construction is said to be characterised by (1) the unmarked NP being the object or goal rather than the actor, (2) the particle *e* marking the non-subject actor NP, and (3) the verbal suffix placing the whole in the perfective aspect stressing the totality of the action (Milner 1974).

This description has much to recommend it, and is, I believe, quite accurate in its assessment of the unmarked NP as the one in focus, and that the particle *e* introduces a non-focal actor NP. The question of the verbal suffixes and their relationship to the nominal particles,



however, may be masked by their consideration as aspectual alone. Milner's work on these various aspects of the Samoan syntactic system has laid the groundwork for this attempt to view it in a Focus framework. It is believed that by examining verbal suffixes and nominal particles as part of an integrated syntactic system, the Focus System, the inconsistencies in earlier analyses may be reduced.

## ANALYSIS

The term focus, as used here, is a sentence level, overt marking system whereby the predicate undergoes derivational affixation to specify one of a restricted number of possible semantic aspects of the happening (typically: actor/agent, goal/object, temporal/spatial specificity, or instrument/motivation). Strictly speaking, focus is neither topicalisation nor emphasis, but an equational relationship between the verb phrase and one of its associated noun phrases. NP's may be thought of as topicalised or emphasised through such devices as preposing, whether or not the sentence contains an NP which is equated with the verb. Conversely, if the in-focus noun phrase does occur it need not necessarily be the NP marked for emphasis by some other device.

To apply a focus analysis I begin, as have others, with the assumption that there are two types of sentences in Samoan: those with suffixed verbs, and those without. I borrow the suggestion that the in-focus NP is unmarked, and that the un-focused actor NP is marked by the particle *e*. This analysis differs in its consideration of the verbal suffix as an integral part of the Focus System, whatever its aspectual implications, and does not consider the particle *i* to mark the object or goal NP, but rather any non-focal, non-actor noun phrase. These assumptions may be represented by the following schema:

Nominal Particle	Abbreviation	Description
'o	CM=	Equational construction marker, marks the focused NP when preposed, and is the appositional marker between two noun phrases.
e	CM≠A	Non-equational actor construction marker, marks the non-focal actor.
i, 'i	CM≠	Non-equational construction marker, marks non-focal, non-actor noun phrases.

With the above assumptions, let us examine the sentence below.

- (5) na tā-ia le tama e le toea'ina 'i le lā'au.  
 t hit-AF the child CM≠A the old man CM≠ the stick

'The old man hit the child with the stick' OR 'The child was hit by the old man with the stick' OR 'The child was hit with the stick by the old man.'

In (5) the verbal suffix<sup>2</sup> indicates that the sentence is non-actor focus (-AF), '*the child*' is unmarked, and so equated with the verb as the focal NP, '*the old man*' is the non-focal actor, and '*the stick*' is both non-focal and non-actor. Activity and passivity in the Indo-European sense must be left to the Indoeuropean speaker to interpret, since syntactically the pertinent information is only that '*the child*' = '*the hit*'. In fact, it may be found through further investigation that '*the old man*' is no more a part of the kernel of sentence (5) than '*the stick*', both NP's being of a commentary nature.

It should also be noted that the problem of glossing the prepositions is a moot one, as the English equivalent must be matched to the environment rather than to a particular particle. This obviates the multiple, homophonous elements required by earlier models.

The two sentence types under consideration here are, therefore, actor-focus (+AF) and non-actor focus (-AF). The nominal particles simply identify a noun phrase as +focus or -focus and +actor or -actor.

To return to the so-called 'passive form, but active meaning' sentences such as (2a), it may be noted that from a basic sentence type of (1) or (2) it is the actor which is pronominalised and preposed in (1a) and (2a) regardless of the focus. This is true for the class of unsuffixed 'passive' verbs such as *fasi* in (4) as well. As Milner (1962:154) pointed out *na ia fasi* is not a transformation of *na fasi e le tama* ('*The child was beaten*'), but of *na fasi e le tama* ('*The child beat something*') as one would expect if pronominalisation and preposition is restricted to the actor NP.

Concerning the sentences in which the verb is not suffixed, although they are presumably non-actor focus, the evidence is inconclusive. At this stage of the analysis the meaning of too few of these sentences is known with enough certainty to posit a strongly supported argument. It may be temporarily sufficient to posit a class of verbs which are inherently non-actor oriented, and consequently require no suffix. A more important and revealing question, however, is raised by sentences of the type shown in example (3) in which the actor is not marked with the non-focal *e* but with *i*. This rather puzzling event is clarified if the researcher does not see 'actor' through his own Indoeuropean lenses, but is willing to allow the semantic category of the NP to be defined by the Samoan syntax. Obviously, in examples such as (3) the non-focal NP under examination is placed in the same category as the 'instrument' in (5), the 'referent' in (6), and the 'direct object' in (1). It has not been placed with the non-focal actors in the same examples.

- (6) 'ua moe le tama i lona fale.  
 t sleep the child CM~~≠~~ her house  
*'The child is asleep in her house.'*

The assertion that the 'o particle marks the NP which is equated with the verb has been somewhat neglected, but may be summarised as follows.

Evidence for the interpretation of the particle 'o as the equational construction marker comes not only from an examination of sentences such as (7), but is comparative as well.

- (7) 'o le la'au ua fa'apa'ū-ina i lalo e le tama i le to'i.  
 CM= the tree t cut -AF CM~~≠~~ down CM~~≠~~A the child CM~~≠~~ the are  
*'The tree was cut down by the boy with the are.'*

In (7) it can be seen that the in-focus NP, *'the tree'*, has been preposed and is introduced by the particle 'o. From examples such as this it is suggested that the CM= is preserved to introduce a preposed, in-focus NP, and may be deleted when the focal NP follows the verb.

- (8) 'o a'u 'o le ta'ita'i.  
 CM= I CM= the leader  
*'I am the leader.'*

In addition, it is seen that the particle 'o may be used appositionally. Through a comparison with a western Austronesian language such as Paiwan, which has a more symmetrical and presumably more conservative focus system, we see that the equational construction marker in that language, a, may also be used appositionally (Ferrell 1974).

- (9) k -m- an a vavaian a aʔak.  
 eat (+AF) CM= woman CM= child  
*'The female child is eating.'*

Thus, in both Samoan and Paiwan it is the CM= which indicates not only an equational relationship between the verb phrase and in-focus NP, but between two noun phrases as well. This apparent structural similarity in the construction marking system, and by extension in the focus system itself, may indicate more syntactic conservatism in the Austronesian language family than is at present assumed.

## CONCLUSIONS

In summary, it has been attempted here to show that previous descriptions of the role of nominal particles in Samoan have not been adequate, and that their interpretation as part of a larger syntactic system, the focus system, is more economical and has the potential for greater explanatory power. Although not sufficiently pursued in this

short paper, the analysis has also suggested that certain semantic categories such as active/passive, subject, object, etc. are conceptually different in Samoan and English, and a more detailed investigation of the syntactic situation may reveal insights not possible through the use of Indoeuropean grammatical models. Finally, it has suggested that more structural similarity may exist between the languages of Polynesia and western Austronesia than has previously been thought. As an additional example, the Samoan focus system seems to bear a strong resemblance in its tendency to reduce a four focus system as is found in the west to +AF and -AF only to the focus system of Indonesian (Bahasa Indonesia). Unlike Indonesian, however, which has lost its focus related construction markers (Thomas 1979), Samoan has remained conservative in its preservation of the focus role of its nominal particles.

It is hoped that further, more detailed study of the syntactic systems of Samoan and other Austronesian languages will shed light on the hypotheses introduced here, and yield further insights into the intrafamily structural relationships of Austronesian and the nature of structural change.

#### N O T E S

1. It should be noted, however, that Marsack (1962:72) provides us with the sentence *ua fasia le pua'a e Ioane* in which the 'passive' form of the verb is used. This particular sentence may, therefore, be poorly chosen to illustrate the class of unmarked verbs.
2. Only Milner's class I suffixes will be discussed here, so that for the purpose of this paper they will be considered to mark non-actor focus. This assumption may have to be modified with the examination of class II in more detail, as it may be found that some suffixes of this class also carry focus responsibilities.

THE SAMOAN CONNECTION

BIBLIOGRAPHY

CHURCHWARD, S.

- 1926      *A New Samoan Grammar*. Melbourne: Methodist Church of Australasia.

FERRELL, R.J.

- 1974      Construction Markers and Subgrouping of Formosan Languages. Paper presented at the First International Conference on Comparative Austronesian Linguistics, Honolulu, Jan. 2-7.

MARSACK, C.C.

- 1962      *Teach Yourself Samoan*. London: English Universities Press.

MILNER, G.B.

- 1962      'Active, Passive, or Perfective in Samoan: A Fresh Appraisal of the Problem'. *Journal of the Polynesian Society* 71/2: 151-61.
- 1974      'It is Aspect (not Voice) which is Marked in Samoan'. *Oceanic Linguistics* 12:621-39.

THOMAS, M.R.

- 1979      'Verbal Affixes and Focus in Bahasa Indonesia'. In: P.B. Naylor, ed. *Austronesian Studies: Papers from the Second Eastern Conference on Austronesian Languages*. Ann Arbor, in press.



PERCEPTUAL DIMENSIONS OF CANTONESE TONES:  
A MULTIDIMENSIONAL SCALING REANALYSIS OF  
FOK'S TONE CONFUSION DATA

JACK GANDOUR

1. INTRODUCTION

It is a common assumption in research on speech perception that a listener's internal representation of the speech signal is organised, at least in part, in terms of phonetic and/or phonological features employed in linguistic analysis. This process of speech perception has been summarised by Studdert-Kennedy (1975:253) as follows:

In short, perception entails the analysis of the acoustic syllable, by means of its acoustic features, into the abstract perceptual structure of features and phonemes that characterize the morpheme.

Many investigators have attempted to determine the number and nature of these perceptual dimensions or features that listeners put together in the identification of speech sounds - consonants (Singh 1975), vowels (Terbeek 1977) and tones (Gandour and Harshman 1978). This paper is an attempt to discover the dimensions or features underlying the perception of Cantonese tones.

Phonological descriptions of Cantonese list six contrastive tones, that may generally be described as (1) high falling, (2) high rising, (3) mid level, (4) low falling, (5) low rising and (6) low level. Compare the following proposed representations of the Cantonese tones, in Chao (1930) tone-number notation:

---

\* The research for this paper was supported in part by a post-doctoral fellowship (1976-1977) at Bell Laboratories, Murray Hill, N.J. I want to especially thank Osamu Fujimura of Bell Laboratories for making his research facilities available.

	1	2	3	4	5	6
Chao (1947)	53	35	33	21	23	22
Kao (1971)	53	35	33	21	23	22
Hashimoto (1972)	53	35	44	21	24	33
Vance (1976)	55	35	33	11	13	22

Of these impressionistic accounts, Hashimoto's (1972) measurements of actual fundamental frequency contours on citation forms appear to reflect the Chao (1947) and Kao (1971) tone descriptions.

Among the three earlier experimental investigations of the perception of Cantonese tones (Fok 1974; Vance 1976, 1977), the Fok study serves as the point of departure for present study of perceptual dimensions of tones. The patterns of confusions (or misidentifications) in her listening identification tests suggest that Cantonese speakers perceive separate pieces of various tonal patterns in making their identifications. If they had perceived each tone as a unitary whole, then on making a mistake, they should have been as likely to guess one tone as any other. But this did not happen; instead, two tones were most likely to be confused if they were similar in their fundamental frequency patterns, and least likely to be confused if their fundamental frequency patterns were highly dissimilar. All this suggests that Cantonese tones are perceived in terms of separate features or dimensions relatively independent of each other.

The present study uses a multidimensional scaling model of perception to investigate the number and nature of features underlying the patterns of confusions among the six Cantonese tones. The dimensions extracted from this reanalysis of the Fok data are evaluated in terms of their perceptual and linguistic plausibility, and in terms of their implications for a more general model of speech perception.

## 2. METHOD

### 2.1. THE INDSCAL MODEL

The output of multidimensional scaling procedures consists of a single map, or configuration, of points - one point for each stimulus. Distances between points reflect the relative similarities among objects; that is, objects which the data indicate to be more similar are in general closer to each other in the map than are less similar pairs.

In many applications of multidimensional scaling in the behavioural sciences, the similarities are obtained from several different subjects, or from the same subjects on different occasions or under different experimental conditions. Recently, a new method was developed by Carroll



and Chang (1970), and implemented in a computer program called INDSCAL (for INDividual Differences SCALing), that determines the common dimensions underlying the similarities data from different subjects or other kinds of data sources, and further determines the relative importance or weight of each dimension to every subject.

The input to INDSCAL consists of many different matrices of similarities or dissimilarities, all pertaining to the same stimulus objects. Each matrix typically comes from one person, but it is also possible for it to be associated with one of several different experimental conditions, measures of similarity, time periods, or locations. As in other multidimensional scaling procedures, the output from INDSCAL includes a map in which each point represents one stimulus object (referred to as the group stimulus space), but unlike other multidimensional scaling procedures, the INDSCAL output also includes a set of dimension weights for each subject (or some other data source) which shows the relative importance of each stimulus dimension to him. Subject weights may be plotted in a map in which each point represents one subject (referred to as the subject space).

In INDSCAL, as in other methods for multidimensional scaling, experimentation is required to determine the number of dimensions that are needed. For any specified dimensionality INDSCAL determines the stimulus co-ordinates, the subject weights, and the unique orientation of axes that account for the maximum total variance in the similarities data from all subjects. The distances between the stimulus objects in some latent psychological space depend on the subjects' dimension weights as well as on the stimulus co-ordinates. The program finds the particular orientation of axes that maximises the goodness-of-fit measure; in most cases, these axes or dimensions can be interpreted without notation. The unrotated dimensions have a special status in INDSCAL, and might be assumed to correspond to fundamental psychological processes that have different saliences for different individuals or under different experimental conditions.

## 2.2. FOK'S (1974) DATA ON PERCEPTUAL CONFUSIONS AMONG CANTONESE TONES

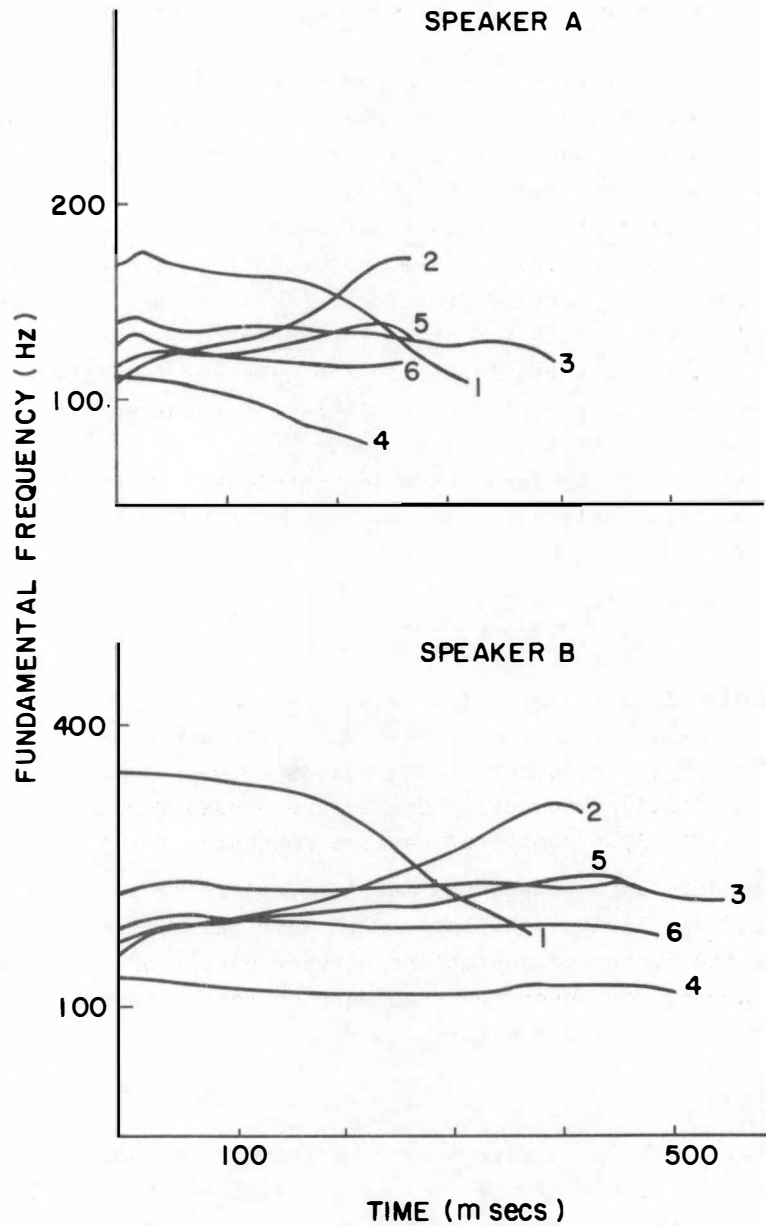
The data for this study are from an experimental investigation of perceptual confusions among Cantonese tones under different experimental conditions (Fok 1974). In one of the experiments, a male speaker (Speaker A) read the following set of words: /fu<sup>1</sup>/ 'man', /fu<sup>2</sup>/ 'bitter', /fu<sup>3</sup>/ 'richness', /fu<sup>4</sup>/ 'to help', /fu<sup>5</sup>/ 'woman' and /fu<sup>6</sup>/ 'father' from long randomised stimulus lists. Under the first experimental condition, the stimuli were simply natural speech versions of

the above set of words, produced at normal tempo with neutral mode of expression. Under the second experimental condition, the stimuli consisted of the natural larynx tones associated with this set of words. These larynx tones were obtained from direct recordings of changes in electrical impedance in the region of the larynx which occur during speech production. Under the third experimental condition, the stimuli consisted of low-pass filtered synthetic versions of these larynx tones. For each of these experimental conditions, the subjects were asked to identify the tone of the stimulus items by circling one of the words in the above set written in Chinese characters. The same experiment was repeated with a female speaker (Speaker B) reading the stimulus set. For detailed discussion of experimental method and procedure, see Fok 1974.

Figure 1 presents the fundamental frequency curves from which the various tone stimulus sets were constructed for both speaker A and speaker B. These two sets of stimuli represent two different sets of stimulus objects, and consequently must be treated separately for the purposes of a multidimensional scaling analysis.

FIGURE 1

Tracings of fundamental frequency curves for the six Cantonese tones (adapted from Fok 1974, by permission of Centre of Asian Studies, University of Hong Kong).



In a comparison of the two stimulus sets, it is to be noted that the speaker A tones generally display shorter duration and a narrower range of fundamental frequency, that the fundamental frequency interval between speaker A tone (1) and his remaining tones is much narrower than the difference in pitch height between speaker B tone (1) and her remaining tones, and that the fundamental frequency shape of speaker A tone (4) displays more of a falling contour than its speaker B counterpart.

A total of 511 subjects participated in the experiments, all native speakers of the Cantonese spoken in Hong Kong. There were three principal groups: 139 secondary school boys, 157 secondary school girls and 215 post-secondary school students.

Eighteen matrices of stimulus-response confusions, or errors of identification (the diagonal values indicated frequencies of correct identifications) were derived from the subjects' judgements, one for each condition for each of the three subject groups. For the speaker A stimulus set, there were nine confusion matrices (3 subject groups x 3 experimental conditions); and similarly, for the speaker B stimulus set, there were nine confusion matrices.

The entire set of eighteen confusion matrices from the Fok study was converted into similarity matrices by using the following equation (Shepard 1972):

$$\text{Sim}(i,j) = \frac{f(i,j) + f(j,i)}{f(i,i) + f(j,j)}$$

where  $\text{Sim}(i,j)$  = similarity between i and j

$f(i,j)$  = number of confusions between i and j

$f(j,i)$  = number of confusions between j and i

$f(i,i)$  = number of correct responses for i

$f(j,j)$  = number of correct responses for j

This procedure symmetrised the matrices, yielding a single value to represent the similarity associated with each pair of stimulus tones. By dividing the number of confusions between a pair of tones by the number of correct responses for that pair of tones removes the effect of bias from the resulting values.

### 2.3. ANALYSIS

Two separate INDSCAL analysis were performed - one on the speaker A stimulus set, the other on the speaker B stimulus set. The input to the INDSCAL analysis of the speaker A stimulus set consisted of nine symmetric (6 stimulus tones x 6 stimulus tones) similarity matrices.

The nine matrices represented perceptual judgements of the Cantonese tones for each of the three subject groups (secondary-level male students, secondary-level female students, post-secondary-level male and female students) across three different experimental conditions (natural speech tones, natural larynx tones, synthetic larynx tones). Each matrix contained similarity estimates for each pair of Cantonese tones for a single condition for a single subject group. The input to the INDSCAL analysis of the Cantonese tones of speaker B was similarly constructed. INDSCAL analyses of these similarity matrices were performed at several dimensionalities in order to determine the appropriate number of dimensions underlying the subjects' perceptual judgements.

### 3. RESULTS AND DISCUSSION

#### 3.1. NUMBER OF DIMENSIONS

For INDSCAL analyses of both the speaker A and speaker B stimulus sets, three dimensions were found to provide the best representation of the perceptual structures underlying the subjects' similarities data. For the tones of speaker A, three dimensions accounted for approximately 96% of the total variance, roughly 22% more variance than two dimensions and only 2% less variance than four dimensions; for the tones of speaker B, three dimensions accounted for approximately 94% of the total variance, about 19% more variance than two dimensions and only 14% less variance than four dimensions. These results provide evidence that the three dimensions are both necessary and sufficient to characterise the subjects' patterns of confusions among the Cantonese tones.

#### 3.2. INTERPRETATION OF DIMENSIONS: SPEAKER A

Plots of the first and second dimensions, and the second and third dimensions from the 3-dimensional INDSCAL group stimulus space of the Cantonese tones for speaker A are shown in the upper-half of Figure 2 (cf. Figure 1).

The first dimension is interpreted as CONTOUR. This dimension separates those tones that have a relatively steady fundamental frequency trajectory (3, 6) from those tones that show considerable changes in fundamental frequency (2, 4, 5). The position of tone (1) would appear to be inconsistent with this interpretation until one takes into account the variant shapes this tone assumes in different phonetic contexts. Tone (1), the high falling tone, becomes high level before another high falling or high level tone (Chao 1947, Huang 1965, Kao 1971), or in Chao tone-letter notation, 53 → 55/ —  $\begin{Bmatrix} 53 \\ 55 \end{Bmatrix}$ . This tone

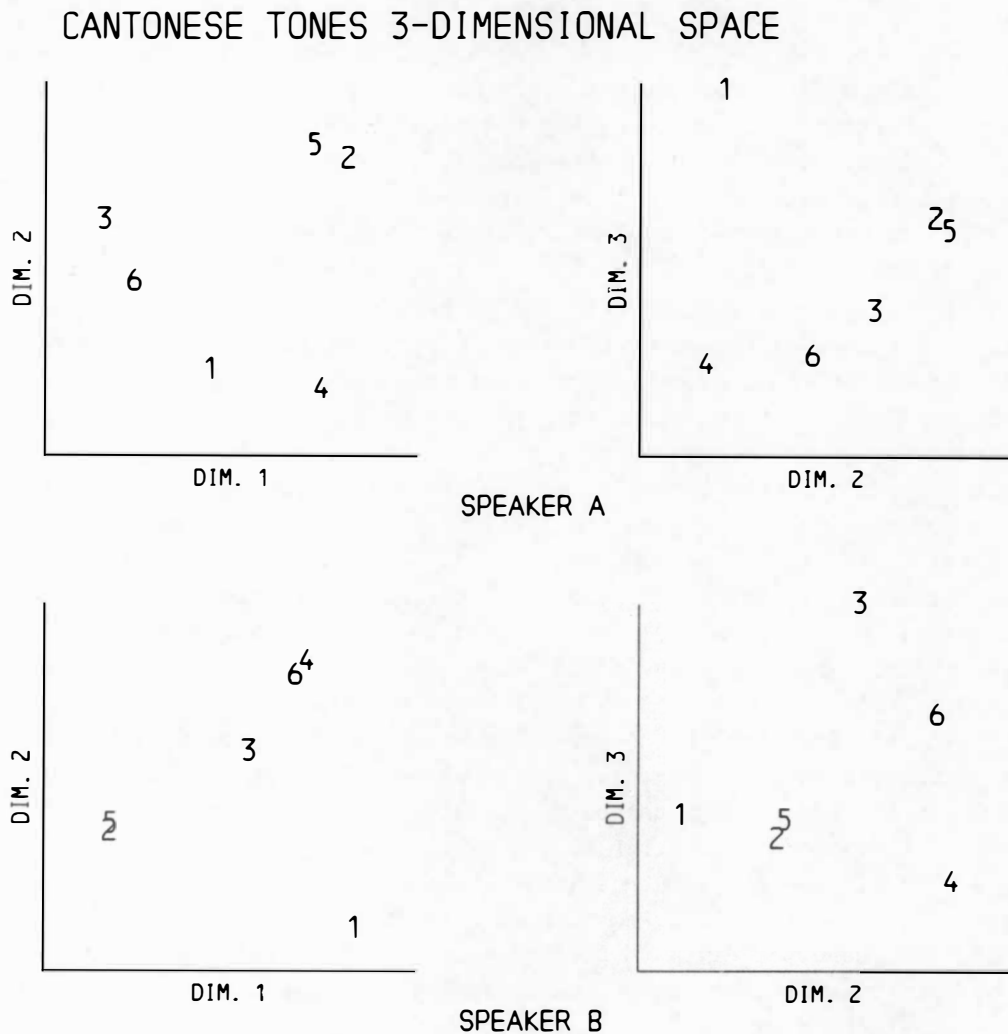
sandhi rule thus establishes a phonological relationship between a contour-shaped and flat-shaped variant of tone (1). It is this phonological relationship that accounts for the intermediate position of tone (1) on this first dimension. Despite the fact that the stimulus version of tone (1) displayed a high falling fundamental frequency pattern, the high level fundamental frequency pattern of the sandhi variant, not present in the stimulus set actually used in the Fok experiment, influenced the subjects' perception of the Cantonese tones. This result is consistent with the findings of the Vance (1977) perceptual study of the six Cantonese tones. The synthetic stimulus tones in his experiment most consistently identified as tone (1) were either high falling or high level.

The second dimension is interpretively labelled DIRECTION. This dimension places the two rising tones (2, 5) and two falling tones (1, 4) at opposite ends of this axis, and the two level tones (3, 6) in the middle. The label HEIGHT is assigned to the third dimension since the order and position of the tones on this dimension appear to be determined by average pitch height. The fact that tone (1) and tone (4) are at opposite ends of the axis as well as the fact that tone (3) falls somewhat near the middle lends support to this interpretation. The position of the tones on the third dimension indicate more crowding in the lower end of the fundamental frequency range. This result is in agreement with earlier impressionistic judgements of the Cantonese tones (Chao 1946, Vance 1977).

In this 3-dimensional INDSCAL group stimulus space, the first, second and third dimensions account for 38%, 36% and 23% of the total variance, respectively. The proportion of variance accounted for by each of the three dimensions indicates that the first and second dimensions are almost identical in saliency while the third dimension is a little weaker than the first two dimensions. The saliency of the first and second dimensions clearly reinforces the notion that, underlying the perceptual judgements of the six contrastive tones of Cantonese, speakers make decisions about whether the tone is level or contour in shape and whether the tone is level, rising or falling in direction.

FIGURE 2

Dimensions 1 and 2 (upper left, lower left) and dimensions 2 and 3 (upper right, lower right) of the group stimulus space from three-dimensional INDSCAL analyses of the Fok tone confusability data for speaker A and speaker B, respectively.



### 3.3. INTERPRETATION OF DIMENSIONS: SPEAKER B

Plots of the first and second dimensions, and the second and third dimensions from the 3-dimensional INDSCAL group stimulus space of the Cantonese tones for speaker B are shown in the lower half of Figure 2 (cf. Figure 1).

The first dimension is interpretively labelled DIRECTION. This dimension places tone (1) and tones (2, 5) at opposite ends of the axis, with tones (3, 4, 6) in the middle. It is nearly identical to the second dimension for the speaker A perceptual space except for the position of tone (4). In the speaker B perceptual space, tone (4) clusters with tones (3, 6); in the speaker A perceptual space, tone (4) clusters with tone (1). This difference in the positioning of tone (4) on the DIRECTION dimension can be understood in terms of the actual shape of tone (4) in the speaker A stimulus set and the speaker B stimulus set, respectively. In the former, tone (4) displays a falling fundamental frequency pattern; in the latter, tone (4) displays a relatively level fundamental frequency pattern. Such realignment of different stimulus tones along what appears to be the same dimension provides further confirmation of its psychological reality.

The second dimension, similar to the first dimension of the speaker A perceptual space, separates the level tones from the contour tones - this is, tones (3, 6, 4) from tones (1, 2, 5). This dimension is labelled CONTOUR. Unlike the first dimension in the speaker A perceptual space, however, tone (1) is not perceived to be similar to the level tones, perhaps because of the much wider pitch interval between tone (1) and tones (3, 4, 6) in the speaker B stimulus set.

The third dimension in the speaker B group stimulus space, while the most difficult to interpret, can be interpreted as a HEIGHT dimension. The order and position of tones (3, 6, 4) along this dimension closely correspond to the actual pitch height of the stimulus versions of these tones. Thus, tone (6) is perceived to be more similar to tone (3) than tone (4). The problem lies with the contour tones (1, 2, 5). These three tones were perceived to be virtually identical along this dimension. Why they cluster toward the lower end of the fundamental frequency range remains unclear.

In this 3-dimensional INDSCAL group stimulus space for the Cantonese tones of speaker B, the first, second and third dimensions account for 40%, 34% and 20% of the approximate total variance, respectively. For both the speaker A and speaker B stimulus sets, the CONTOUR and DIRECTION dimensions together accounted for 74% of the variance. Their relative importance, however, varies across the stimulus sets, with the



DIRECTION dimension being more predominant in the perceptual judgements of the Cantonese tones of speaker B. The third dimension, HEIGHT, turns out to be the weakest of the three dimensions in this experimental task, for both the speaker A and speaker B sets of Cantonese tones. The emergence of three interpretable dimensions in the two separate INDSCAL analyses, similar in the order and position of the tones along the dimensions and plausible from both a perceptual and linguistic standpoint, confirms the validity of the multidimensional scaling analysis. It is surely not accidental that the 'same' three dimensions underlie these two slightly different sets of Cantonese tones, and that the relative importance of these three dimensions to each other depends on the exact composition of the stimulus tones.

#### 3.4. DIMENSION WEIGHTS FOR EXPERIMENTAL CONDITIONS

In general, there was little systematic variation in the patterns of dimension weights for the three different kinds of stimuli - natural speech, natural larynx tones and synthetic larynx tones. This would seem to indicate that pitch variation is the principal cue for distinguishing the six Cantonese tones. In other experimental investigations of Cantonese tones, Fok (1974) and Vance (1976) have shown that concomitant differences in syllable duration are not important for Cantonese tonal distinctions.

For the speaker A set of Cantonese tones, however, weights on the CONTOUR dimension are highest under the natural speech condition; for the speaker B set of Cantonese tones, weights on the HEIGHT dimension are highest under the natural speech condition. This perhaps reflects the differences in degree of movement in fundamental frequency patterns between the two stimulus sets. Under the natural speech condition, subjects optimally directed their attention to these general properties of the two sets of tones. While this is not a wholly satisfactory explanation, nevertheless, these results do show that the dimensions obtained and their relative importance may depend on the experimental conditions as well as the stimuli and subjects. In the Fok study, no differences in the perception of Cantonese tones could be attributed to the three different subject groups. This result is demonstrated in the INDSCAL analyses in the lack of differences in the patterns of dimension weights across subject groups.

#### 4. FURTHER DISCUSSION

The particular three dimensions that emerged from the INDSCAL analyses of the Cantonese tone confusions - CONTOUR, DIRECTION and

HEIGHT - are meaningfully interpretable in both perceptual and linguistic (cf. Wang's (1967) proposed set of phonological features of tone) terms. This is important for providing us a better understanding of how Cantonese speakers construct an internal representation of the tonal patterns. The fact that we obtained perceptually and linguistically plausible dimensions for Cantonese tones takes on even greater significance when compared to the results of other multidimensional scaling investigations of tone perception. Gandour and Harshman (1978), in their cross-language study of tone perception, which included two typologically and genetically unrelated tone languages (Thai, the national language of Thailand, and Yoruba, a language spoken primarily in Nigeria), also found dimensions that could be similarly labelled. In another multidimensional scaling analysis of the perception of tones in Yoruba, Hombert (1976) found dimensions that could be related to the direction of movement in fundamental frequency, and the distinction between level and contour fundamental frequency shapes. The present study as well as these two earlier ones employed different stimuli, subjects and experimental tasks; yet the dimensions extracted are similar enough to suggest that such dimensions or features must, indeed, be psychologically real, and part of the universal set of phonetic/phonological features underlying the perception of tone. The precise number and nature of these features, of course, is a subject for future experimental investigation.

Linguists are not only interested in the construction of a universal set of phonetic/phonological features, but they also wish to know to what extent these features are utilised in particular languages. Or to put it in a different perspective, how does the language background of a listener affect his perception of speech sounds; in the context of this study, does the structure of Cantonese influence his perception of the six contrastive tones? Although our data lacks a control group for comparison, the position of tone (1) on the CONTOUR dimension in the speaker A group stimulus space strongly suggests that phonological rules may influence a person's perception of speech sounds. This effect of particular linguistic experience is also shown in the Gandour and Harshman (1978) study, where differences in the composition of tonal inventories, or the lack thereof, is reflected in differential emphasis placed on selected dimensions by speakers of Thai, Yoruba and English. Such data point to the need for a model of speech perception that incorporates higher-level linguistic information into the perceptual processing of speech signals.

And finally, the application of INDSCAL has proven to be a very useful tool for learning a great deal about both the stimuli and the subjects under investigation, for discovering the underlying dimensions of a multidimensional perceptual space, and for confirming and improving on specific hypotheses concerning human perception. These findings can now be incorporated into the design of other laboratory experiments dealing with the processing of tone as well as other speech sounds.

## 5. SUMMARY

A multidimensional scaling reanalysis of the Fok (1974) tone confusion data from Cantonese revealed three underlying perceptual dimensions that were interpretively labelled CONTOUR, DIRECTION and HEIGHT, respectively. The influence of a Cantonese tone sandhi rule on a listener's tonal perception was evident in the position of the high-falling tone (1) on the CONTOUR dimension. The weights of these dimensions were not found to vary much across the different kinds of tonal stimuli - natural speech tones, natural larynx tones and synthetic larynx tones. These dimensions extracted from the patterns of tonal confusions in Cantonese were found to bear close resemblance to perceptual dimensions of tone that have emerged in other multidimensional scaling investigations.

BIBLIOGRAPHY

CARROLL, J.D. and J.J. CHANG

- 1970 'Analysis of Individual Differences in Multidimensional Scaling via an n-way Generalization of "Eckart-Young" Decomposition'. *Psychometrika* 35/3:283-319.

CHAO, Y.R.

- 1930 'A System of "Tone Letters"'. *La Maître Phonétique* 32: 24-7.
- 1947 *Cantonese Primer*. Cambridge, Mass.: Harvard University Press.

FOK CHAN YUEN-YUEN

- 1974 *A Perceptual Study of Tones in Cantonese*. Centre of Asian Studies Occasional Papers and Monographs, 18. Hong Kong: Centre of Asian Studies, University of Hong Kong.

GANDOUR, J.T. and R.A. HARSHMAN

- 1978 'Crosslanguage Differences in Tone Perception: A Multi-dimensional Scaling Investigation'. *Language and Speech* 21/1:1-33.

HASHIMOTO, O-K.Y.

- 1972 *Studies in Yüe Dialects 1: Phonology of Cantonese*. London: Cambridge University Press.

HOMBERT, J.M.

- 1976 'Perception of Tones of Bisyllabic Nouns in Yoruba'. *Studies in African Linguistics, Supplement 6*, 109-21.

HUANG, P.

- 1965 *Cantonese Sounds and Tones*. New Haven, Conn.: Yale University, for Eastern Publications.

KAO, D.L.

- 1971 *Structure of the Syllable in Cantonese*. The Hague: Mouton.

SHEPARD, R.N.

- 1972 'Psychological Representation of Speech Sounds'. In: E.E. David and P.B. Denes, eds *Human Communication: A Unified View*, 67-113. New York: McGraw-Hill.

SINGH, S.

- 1975 'Distinctive Features: A Measurement of Consonant Perception'. In: S. Singh, ed. *Measurement Procedures in Speech, Hearing and Language*, 93-155. Baltimore: University Park Press.

STUDDERT-KENNEDY, M.

- 1975 'Speech Perception'. In: N.J. Lass, ed. *Contemporary Issues in Experimental Phonetics*, 243-93. Springfield, Illinois: Charles C. Thomas.

TERBEEK, D.

- 1977 'A Crosslanguage Multidimensional Study of Vowel Perception'. *Working Papers in Phonetics, University of California, Los Angeles* 37.

VANCE, T.J.

- 1976 'An Experimental Investigation of Tone and Intonation in Cantonese'. *Phonetica* 33/5:368-92.
- 1977 'Tonal Distinctions in Cantonese'. *Phonetica* 34/2:93-107.

WANG, W.S-Y.

- 1967 'Phonological Features of Tone'. *International Journal of American Linguistics* 33/2:93-105.



## REVIEW ARTICLE:

ROBERT K. HEADLEY JR, KYLIN CHHOR,  
LAM KHENG LIM, LIM HAK KHEANG, CHEN CHUN:  
CAMBODIAN-ENGLISH DICTIONARY

PHILIP N. JENNER

Robert K. Headley Jr, Kylin Chhor, Lam Kheng Lim, Lim Hak Kheang, Chen Chun, *Cambodian-English Dictionary*. Bureau of Special Research in Modern Languages. Publications in the Languages of Asia, III (Washington, D.C.: The Catholic University of America Press, 1977), Volume I: K - BH, Volume II: M - '. Exclusive Distributor: ISBS, Inc., P.O. Box 555, Forest Grove, OR 97116. U.S. price: \$49.00.

Physically, this is a massive unabridged work in two volumes measuring 21.5 by 28 cm (8½ by 11 inches). Volume I weighs about 2.25 kg (4.9 lbs.), Volume II about 2.75 kg (5.9 lbs.). The reproduction is by photo-offset from camera-ready typescript, and is done on good stock of uncommon whiteness. It is bound with not unattractive casings in navy cloth with gilt lettering on the spines and fronts.

The main body of the dictionary (1-1495) is preceded by a Preface (v), Acknowledgements (vii), a table of Contents (ix), a list of Plates (xi), an Introduction (xiii-xxii), a list of Abbreviations (xxiii-xxiv), and a Select Bibliography (xxv-xxvii) of forty titles, "all of the works systematically consulted for vocabulary items or other information during the preparation of the dictionary". The body of the work is followed by thirteen plates, each containing one or more line-drawings giving Khmer terms for a considerable number of cultural and other objects - mostly household articles, house parts and types, agricultural and fishing implements, loom parts, games, boat types and musical instruments - and a few plants.

The Introduction comprises two paragraphs on the background of Khmer (xiii), with which is to be included a full-page chart (xiv) showing the classification of selected Mon-Khmer languages. This is followed by sections on the Arrangement of the Dictionary (xiii-xv), Phonology (xv-xvii), Khmer Syntax and Word Classes (xvii-xviii), and the Khmer Script (xix-xxii).

In the dictionary proper pages are laid out in generous double columns. All Khmer forms have been done on the Vari-Typer, the Khmer font for which is well designed and legible except for some of the smaller conjunct symbols. In a few combinations a magnifying glass is needed to make out the *anusvāra*, the breve, and certain of the conjunct vowels. Matter in roman type looks as if it had also been done on the Vari-Typer, though justified right margins are few.

Main entries, in the Khmer character, are followed by phonemic transcriptions. The pronunciation represented is "that of Standard Khmer as spoken by educated persons throughout the Khmer-speaking area" (xv). English meanings are preceded by an italicised abbreviation indicating the wordclass to which the item belongs. These headwords may also contain "alternate pronunciation(s), synonyms ..., usage indicators ..., examples, etymological information, alternate spellings ..., feminine forms [of Indic loans], antonyms ..., or cross references to related or derived forms or to entries which might provide additional information ..." (xv). Main entries so described are followed wherever appropriate by a considerable selection of subentries consisting for the most part of constructs in which the headword is the second as well as the first member. Phonemic transcriptions are not given after these subentries, a circumstance which will inconvenience those who do not yet read well but which represents a major saving of space. The total number of entries and subentries is not stated, but would appear to be close to 90,000.

While the Preface (v) notes that Dr Headley and his team of native-speaking assistants "have tried to utilize all previous Khmer lexical material", the dictionary is based on the fifth edition of the semi-official, two-volume *Vacanānukrama khmēra / Dictionnaire cambodgien* (Phnom-Penh: Institut Bouddhique, 1967, 1968). "Great numbers of additional words, expressions, and examples were extracted from many secondary sources ...", chiefly from Sam Thang, *Vākyaparivattana khmēra-pārāṃṇa / Lexique khmer-français* (Phnom-Penh, 1962) and S. Tandart, *Dictionnaire cambodgien-français* (Phnom-Penh: Albert Portail, 1935). It is a cause for regret that a few items registered in the VK are not listed here. Two examples are *keṣṭ* /kēsəḏəj ~ kaaəsəḏəj/ '(rā)jasabda) head' (cf. Sanskrit keśin) and *kūva* /kōw/, heretofore one of the most



enigmatic words in the older language, which Khin Sok (1979) has recently narrowed down to 'précisément' (e.g. *thñai neḥ kūva /tḥaj néh kòow/ 'this very day'*). One suspects that the basis for the exclusion of these and a good many other items was the circumstance that they are archaic.

This being the first unabridged Khmer-English dictionary by a trained linguist, no one should be surprised that it falls short of perfection. However, its strengths are as many as its weaknesses, though a work of such ambitious scope may be particularly liable to faultfinding. Its most conspicuous strengths are its fullness and the evident pains to which Dr Headley and his team have gone to ensure accuracy of detail. Its chief weaknesses, aside from a pervading unevenness of quality, are four: (1) an unjustified dependence on the *VK*, (2) insufficient attention to the older literature as a source of lexical data, (3) several inconveniences in the phonemic transcription, and (4) several inconveniences in the alphabetisation. The first of these is only to be expected in the present state of Khmer studies; a full translation of the *VK* alone would have been a welcome undertaking. The second weakness, again because of the state of our knowledge, is entirely predictable and would have required a formidable effort to overcome. The other two weaknesses, less basic, are of more immediate interest.

Headley follows Huffman (1970:24-8) in excluding register from his analysis of the vowel inventory. This is a perfectly valid approach, though it seems to this reviewer one which should be carefully explained if it is not to entail features which are liable, in a work of this importance, to lead the reader astray and, worst of all, to be misunderstood by persons who may have to cite forms from the dictionary without any knowledge of Khmer.

In Headley's analysis the vowel phonemes are neatly divided into simple and complex (falling diphthongs).

In the simple nuclei nine short phonemes are recognised. These are written as unmarked units, and include, as in my own analysis, a sub-phonemic /ɔ/ found in High Register presyllables. The eleven long nuclei, marked by gemination, include an /ee/ and an /eɛ/ representing respectively the High and Low Register realisations of orthographic *e*. Thus *jera* [čeː: ~ če:] '*to insult*' is Mrs Jacob's (1968:3-24; 1974) phonological *cè:(r)*, Huffman's and Headley's simple /cee/, and my /céer/; while *cera* [čeːː ~ če:] '*of long duration*' is Jacob's *ce:(r)*, Huffman's /cei/, Headley's /ceɛ/, and my /cèer/. This much is well and good, especially if it is explained to the reader that the /ei/ and /eɛ/ are phonemic devices marking a relationship with the registral

system or the writing system. In the back rounded mid vowels, however, this graphic diphthongisation has no counterpart. Thus *jora* [čɔːː ~ čɔ:] 'to rise, mount' is Jacob's cò:(r), Huffman's and Headley's /coo/, and my /cóor/, while *cūra* [čɔːː ~ čɔ:] 'mild imperative marker' is Jacob's co:l [sic], Huffman's /cou/, my /còor/, but Headley's /coo/ again. Thus the symmetry which Huffman achieves is lost.

The ten complex nuclei, inexplicably, are not discriminated according to their length. The seven long nuclei include an /ie/ defined as [i:ɛ] (xvi), representing the High Register realisation of orthographic ā, and an /iə/, undefined as to length, representing both registers of orthographic iə. Thus *dāra* [ti:ɛ] 'to dun' is Jacob's tì:ə(r), Huffman's /tiə/, Headley's /tie/, and my /tíiər/, while *diəta* [ti:ət] 'more, still' is Jacob's tìət, Huffman's and Headley's /tiət/, and my /:fiət/ (cf. *tiəma* [ti:əm] 'shop' = Jacob's tiəm, Huffman's and Headley's /tiəm/, my /tìiəm/). The other long diphthongs are /+ə/, /uə/, /ae/ (correctly defined as "Front Low to Mid-Low"), /æ/, and /ao/. The three short diphthongs are /ea/, defined as [ɛa] though written [ɛa]; /oa/, which has two allophones: [oa] ("short mid-high back to low back") before zero, and [ɔæ], written [ɔa] ("short mid-low back to low front") before finals other than /k, ʔ, ŋ, h/. Thus *bā'ka* [pɛək] 'to hang' is Jacob's phonological pɛək, Huffman's /pěəq/, my /peək/, Headley's /peak/; *dā'ta* [tɔ'ət] 'to kick' is Jacob's tɔət, Huffman's /tōət/, my /toət/, Headley's /toat/; *dāmbāra* [tʊm'po:ə] 'page (of book)' is Jacob's tʊmpɔə(r), Huffman's /tumpɔə/ (note the absence of the breve), my /tumpoər/, Headley's /tumpoa/. There is nothing at fault here except the fuzziness of the phonetic definitions. The third and last of the short diphthongs, however, is correctly defined as [uə] but for some unaccountable reason is phonemicised /ɔə/. In reality this is an environmentally conditioned allophone of /u/ and hence another subphonemic feature. Thus *ga'ta* [kuət] 'precisely' is Jacob's phonological kùət, Huffman's /kuət/, Headley's /kɔət/, my /kut/; *la'ka* [lʊək] 'to sell' is Jacob's lùək, Huffman's /lũəq/, Headley's /lɔək/, my /luk/. Since Headley has stated exactly what he means, the most that can be said is that such an unfortunate choice of symbols can only add to the reader's confusion.

In general, the alphabetisation of entries in the dictionary follows the traditional Khmer order as perfected in the VK. However, two innovations have been introduced, no doubts on grounds of logic, neither of which is especially welcome.

In the first place, vowels before *visarga* are grouped together after those with *anusvāra*, so that *kih*, *kuh*, *keh* and *koh* follow *kaḥ*. In the

*VK* *kaḥ* comes at the very end of the series, the other combinations being interspersed with the vowels in question; for example, *kuḥ* comes after *kuhaka* and before *kūka*.

In the second place, forms with initial *a-kāra* (the *Vokalträger*), with and without conjunct vowel symbols, are grouped into one section all the way through *aṃ* and *aḥ* to *'hoḥ*. This section is then followed by separate sections for the other Indic syllabic vowels such as *i-kāra* and *ī-kāra*. In the *VK*, *a-kāra* with the conjunct vowels is mixed in with the other syllabic vowels so that *'ūna*, for example, is followed immediately by *ūna*, which in turn is followed immediately by *'üṇadrū*.

In both of these features Headley's way is obviously better, but is a nuisance.

One decided improvement over the *VK* is the inclusion of the Indic vowels *ṛ*, *ṝ*, *ḷ* and *ḹ* among the vowels. In the *VK* they are grouped, contrary to tradition, at the end of *ra-kāra* and *la-kāra*.

In a work of this quality, finally, we are entitled to expect that Dr Headley would have taken the bull by the horns and decided upon a policy for the treatment of the *repha*, the allograph of *ra-kāra* for syllable-final */-r/* in the *devanāgarī* and other Indic writing systems. The Khmer have usually treated this as a diacritic, and as a consequence it has never had any fixed alphabetic order. Both the *VK* and Headley therefore vacillate in their ordering of forms with which it occurs. In the *VK* *var-ṇaḥ* comes (correctly) between *va-ruṇaḥ* and *va-rman*, but *sar-bejña* comes (incorrectly) before *sa-rabraḥ*. In Headley *dhar-ma* comes (correctly) between *dha-ramāna* and *dhā'ka*, but *'ār-tha* comes (incorrectly) between *'ātmopajīvin* and *'ādara*.

For all of these negative judgements, Dr Headley, his team of assistants, and the Catholic University of America are to be congratulated for their imagination and perseverance in producing a work of this magnitude and this excellence.

BIBLIOGRAPHY

- HUFFMAN, Franklin E., with Chhom-Rak Thong Lambert and Im Proum  
1970 *Cambodian System of Writing and Beginning Reader with Drills and Glossary*. Yale Linguistic Series. New Haven: Yale University Press.
- JACOB, M.  
1968 *Introduction to Cambodian*. London: Oxford University Press.  
1974 *A Concise Cambodian-English Dictionary*. London: Oxford University Press.
- KHIN SOK  
1979 'L'inscription moderne de Ta Prom (Bat1), K.39'. *Mon-Khmer Studies* 8. Forthcoming.
- SAM THANG  
1962 *Vākyaparivattana khmēra-pārāṇḥa / Lexique khmer-français*. Phnom-Penh.
- TANDART, S.  
1935 *Dictionnaire cambodgien-français*. Phnom-Penh: Albert Portail.
- VK  
1967 *Vacanānukrama khmēra / Dictionnaire cambodgien*. Phnom-Penh:  
1968 Institut Bouddhique. 2 vols.